ANTONI SUŁEK

Eksperyment w badaniach społecznych

ANTONI SUŁEK

Eksperyment w badaniach społecznych

Spis treści

OD AUTORA 7
CZĘŚĆ I ISTOTA I WARTOŚĆ EKSPERYMENTU 9
1. POJĘCIE EKSPERYMENTU 9
2. ZASTOSOWANIA I SZCZEGÓLNE ZALETY EKSPERYMENTU 13
3. KRÓTKA HISTORIA EKSPERYMENTU 25
CZĘŚĆ II SCHEMATY I PROCEDURY EKSPERYMENTALNE W BADANIACH SPOŁECZNYCH 31
4. MANIPULACJA EKSPERYMENTALNA 31
5. ZAKŁÓCENIA BODŹCA EKSPERYMENTALNEGO 35
6. SCHEMATY EKSPERYMENTALNE 42
7. EKSPERYMENTY SPOŁECZNE 76
CZĘŚĆ III GRANICE EKSPERYMENTU W BADANIACH SPOŁECZNYCH 89
8. TECHNICZNE GRANICE EKSPERYMENTU 89
9. AKSJOLOGICZNE GRANICE MANIPULACJI EKSPERYMENTALNYCH 98
10. TZW. SZTUCZNOŚĆ SYTUACJI EKSPERYMENTALNYCH JAKO BARIERA EKSPERYMENTU 109
Część IV REAKCJE NA PRZESZKODY. EKSPERYMENTY ZASTĘPCZE NA MODELACH 130
11. POJĘCIE MODELU I EKSPERYMENTU MODELOWEGO . . 131
12. MAŁA GRUPA JAKO MODEL EKSPERYMENTALNY 135
13. EKSPERYMENTY MODELOWE NA ZWIERZĘTACH 143
SŁOWO KOŃCOWE 155
NAJWAŻNIEJSZA LITERATURA 156

Państwowe Wydawnictwo Naukowe
Warszawa 1979

Okładką projektował Andrzej Pilich
Redaktor Maria Dietl
Redaktor techniczny Janusz Malinowski
Korektor Zespół
Tytuł dotowany przez Ministerstwo Nauki, Szkolnictwa Wyższego i Techniki
© Copyright by Państwowe Wydawnictwo Naukowe
Warszawa 1979

ISBN 83-01-00032-5

Printed in Poland
Państwowe Wydawnictwo Naukowe Wydanie pierwsze. Nakład 1800 + 200 egz.
Arkuszy wyd. 16. Arkuszy druk. 22,5
Papier drukowy m/gł kl. IV 65 g 82×104
Oddano do składania we wrześniu 1977 roku
Podpisano do druku w listopadzie 1978 roku
Druk ukończono w marcu 1979 roku
Zakłady Graficzne w Toruniu
Oddział 3/4 w Wąbrzeźnie
Zam. 3960 10.XI.1977. — Z-16
Cena zł 45,—

Pliki przygotowane na podstawie wydania oryginalnego. Numeracja stron i przypisów może być zmieniona w stosunku do oryginału. W przypadku odwołań i cytowań, prosimy o korzystanie z numeracji wydania oryginalnego.

Wszystkie prawa zastrzeżone. Swoboda użytkownika ograniczona do ustawowego zakresu dozwolonego użytku

Język: polski

OD AUTORA

„Eksperymentowanie z życiem — czyż mogło być coś bardziej fascynującego?" — to retoryczne pytanie wkłada Skinner w usta Fraziera, założyciela eksperymentalnej społeczności Walden Two ze swej głośnej powieści pod tym tytułem. Wydało mi się, że coś takiego, jak eksperymentowanie z życiem (społecznym) może być równie pasjonujące dla metodologa.

Eksperyment w badaniach społecznych jest tworem złożonym. Jest on nie tylko metodą badania, lecz także pewną formą działania i zmiany społecznej, obwarowany jest pewnymi wartościami, posiada swoistą wewnętrzną dynamikę psychospołeczną. Zatem, zagadnienie eksperymentu w badaniach społecznych to problem nie tylko metodologiczny w sensie ścisłym, lecz również socjologiczny, aksjologiczny i psychologiczny.

Próba pokazania możliwości i trudności eksperymentu musi się więc odwołać do argumentów czerpanych z wielu różnych dyscyplin.

Zaciekawienie jakimś problemem wystarcza do tego, by się i im zająć — nie wystarcza do tego, by wyniki swej pracy podawać do wiadomości publicznej. Trzeba wpierw uzasadnić wagę zagadnienia.

Eksperyment jest ważną metodą wielu nauk społecznych. Korzysta zeń przede wszystkim socjologia małych grup i psychologia społeczna — zrąb tych nauk ma rodowód eksperymentalny. Bywa on również często stosowany w innych dziedzinach badań społecznych, choćby w socjologii przemysłu czy w pedagogice, W ostatnich latach w niektórych krajach zarysowała się silna tendencja do eksperymentalnego sprawdzania programów zmiany społecznej przed ich wprowadzeniem w życie. Zatem, znajomość metodologii eksperymentu może być przydatna w projektowaniu i interpretacji eksperymentów.

Znajomość ta może przydać się nie tylko tym, którzy chcieliby prowadzić badania eksperymentalne. Zasady analizy przyczynowej wypracowane na gruncie eksperymentu są bardzo użyteczne w badaniu następstw zmian naturalnych, dokonujących się bez udziału badacza, takich jak innowacje czy reformy społeczne. Eksperyment jest narzędziem wymagającym szczególnie głębokiego i intensywnego kontaktu badacza z rzeczywistością, więc i szczególnie intensywne są interakcje społeczne zawiązujące się w trakcie badania między zaangażowanymi weń stronami. Wyniki badań nad społeczną sytuacją eksperymentu mogą pomóc w zrozumieniu i kontroli sytuacji społecznej badania w innych niż eksperymentalne badaniach opartych na komunikowaniu się badacza z badanymi, zwłaszcza w badaniach ankietowych.

Problematyka eksperymentu w badaniach społecznych posiada również pewne odniesienie ogólnometodologiczne; W szczególnie wyraźnej postaci widać tu choćby jak rzeczywistość stawia opór badaniu naukowemu i jak nauka radzi sobie z trudnościami stojącymi na drodze jej rozwoju. Wyniki badań nad eksperymentem w badaniach społecznych są również istotne dla dyskusji nad możliwościami przenoszenia metod nauk przyrodniczych na teren nauk społecznych. Eksperyment jest podstawową metodą przyrodoznawstwa, więc pokazanie możliwości i granic jego stosowania w naukach społecznych dostarcza argumentów, a przynajmniej ilustracji dla tez wypowiadanych w tej dyskusji.

Dla tych powodów eksperymentem w badaniach społecznych zajmować się warto.

Praca, którą przedstawiam ma za zadanie odpowiedzieć na cztery pytania:

1. Na czym polega specyfika eksperymentu i jego wartość dla nauk empirycznych a zwłaszcza dla badań społecznych?
2. Jakie formy przybiera eksperyment w badaniach społecznych?
3. Jakie istnieją ograniczenia stosowania eksperymentu do badania zjawisk społecznych?

4. W jaki sposób nauki społeczne radzą sobie z tymi ograniczeniami?

Praca składa się z czterech części. Zawierają one odpowiedzi na kolejne pytania.

Książkę tę napisałem na kanwie pracy doktorskiej, przygotowanej w Zakładzie Metodologii Badań Socjologicznych Instytutu Socjologii Uniwersytetu Warszawskiego. Wysoko cenię sobie uwagi, jakie poczynili nad tekstem tej pracy jej promotor Profesor Stefan Nowak oraz recenzenci Profesorowie Jan Lutyński i Klemens Szaniawski.

Róży, mojej żonie, winien jestem serdeczną wdzięczność za wyrozumiałość, z jaką znosiła moją tak długą nieobecność myślami w domu.

A.S.

Warszawa — Jelonki w kwietniu 1977 r.

CZĘŚĆ I ISTOTA I WARTOŚĆ EKSPERYMENTU

Metoda eksperymentalna odegrała wybitną rolę w historii nauk przyrodniczych, tak fizykalnych jak i biologicznych. Zajmuje ona też poczesne miejsce w ich współczesnym instrumentarium, dowodem skłonność do utożsamiania ich z naukami eksperymentalnymi. Widocznie eksperyment posiada takie zalety, jakimi nie mogą poszczycić się inne procedury badawcze. Zalety te spowodowały ekspansję eksperymentu na teren niektórych nauk społecznych a zwłaszcza socjologii małych grup i psychologii społecznej. ’

W niniejszej części przedstawimy samo pojęcie eksperymentu — scharakteryzujemy je w opozycji do badań nieeksperymentalnych, czysto obserwacyjnych. Główną uwagę skupimy jednak na funkcjach eksperymentu a zwłaszcza na sprawdzaniu hipotez przyczynowych, ta bowiem funkcja sprawia, że cieszy się on tak wysokim respektem. Naszkicujemy też krótką historię eksperymentu by pokazać drogę, jaką on przebył, zanim zaczął być stosowany w badaniach społecznych.

Analiza eksperymentu jaką dajemy w tej części pracy posiada w znacznym stopniu charakter ogólno- metodologiczny, odnosi się ona również do eksperymentu w naukach przyrodniczych. Jest w tym pewien zamysł, albowiem ujawnione w tych naukach zalety eksperymentu zachęciły wiele innych dyscyplin do sięgnięcia po to narzędzie. Ogólnometodologiczna charakterystyka eksperymentu będzie tłem, na którym wyraźniej widoczna będzie specyfika eksperymentu jako metody badań społecznych.

1. POJĘCIE EKSPERYMENTU

Podstawą nauk empirycznych jest kontakt z empirią czyli doświadczeniem. Twierdzenia o faktach są fundamentem, na którym wznosi się gmach tych nauk. Jego budowa nie polega jednak na zbieraniu surowych faktów i — następnie dopiero — ich interpretacji i analizie. Związki między twierdzeniami o faktach a uogólnieniami, a zwłaszcza teoriami, są o wiele bardziej subtelne. Jednakże fakty są tą instancją, przed którą w ostatecznym rachunku sądzi się hipotezy, generalizacje i teorie.

Wiedzę o faktach zdobywa się za pomocą obserwacji, rozumianej jako postrzeganie kierowane pytaniem Obserwacja wymaga często — a w naukach rozwiniętych jest to zasadą — posługiwania się skomplikowanymi narzędziami, pozwalającym zwiększyć jej zasięg i precyzję. Teleskop, sejsmograf, chronometr, kardioskop, test psychologiczny są przykładami takich narzędzi.

Obserwacja polegać może na rejestrowaniu naturalnych, istniejących niezależnie od woli badacza stanów rzeczy i ich zmian. Aby obserwacja tego rodzaju była możliwa i dokładna, trzeba niejednokrotnie czynnie interweniować w przebieg badanych zjawisk. Nie robi się jednak tego po to, by zbadać skutki interwencji. Nie zmienia się (a przynajmniej nie zamierza się zmieniać) tych aspektów zjawisk, które mają być przedmiotem badania. Chce się tylko uczynić względy badane łatwiej obserwowalnymi. W badaniach mikroskopowych nad strukturą tkanki organicznej barwi się preparat nie po to by zbadać skutki barwienia ani nie po to, by zmienić strukturę, lecz po to, by uczynić ją bardziej wyraźną.

Jednakże obserwowane stany rzeczy i ich zmiany nie zawsze występują niezależnie od woli badacza. Często są one planowo przezeń tworzone a ich wywoływanie jest istotnym składnikiem metody. Metody eksperymentalnej, ta bowiem opiera się na celowym wpływie na naturalne stany i procesy.

Eksperyment jest to powtarzalny zabieg polegający na planowej zmianie przez badacza jednych czynników w badanej sytuacji, przy równoczesnej kontroli innych czynników, podjęty w celu uzyskania w drodze obserwacji odpowiedzi na pytanie o skutki tej zmiany.

Była to definicja[1]. Jej *definiens* zawiera — jak widać — wskazanie celu całego przedsięwzięcia oraz opis szczególnego sposobu jego realizacji. Cel ma charakter poznawczy: zbadanie skutków zmiany. Jest to cel bezpośredni i może on być podporządkowany różnym celom nadrzędnym, zwykle — sprawdzaniu hipotez o związkach przyczynowych między czynnikami zmienianymi a pewnymi innymi. Drogą do tego celu wiodącą jest kontrolowane wywołanie zmiany, której skutków badacz jest ciekawy. Szczególny charakter tej drogi polega na manipulacji a nie na kontroli. Nie oznacza to bynajmniej lekceważenia tej drugiej. Przeciwnie — znaczy to tyle, że w badaniach nieeksperymentalnych zadany cel można osiągać często i bez manipulacji, ale nigdy nie bez kontroli. Niezależnie od tego czy bada się skutki zmiany eksperymentalnej czy naturalnej, zawsze musi się mieć możliwość odróżnienia wpływu zmiany badanej od. wpływów zmian ubocznych.

Podane rozumienie terminu „eksperyment" nie jest — niestety — jedynym w naukach empirycznych a w socjologii w szczególności. Zakres tego terminu bywa nagminnie rozciągany, a to przez rezygnację z któregoś ze wskazanych warunków definiujących.

Tak więc rezygnacja z warunku pierwszego tj. celu poznawczego zrodziła obyczaj nazywania eksperymentami nowatorskich, odbiegających od rutyny przedsięwzięć społecznych, mimo że ich celem nie jest osiąganie celów poznawczych, lecz praktycznych. Owe „eksperymenty" społeczne łączą z eksperymentami w sensie ścisłym tylko elementy planowania i nowatorstwa oraz niepewność wyników. Różni zaś — wywodzący się z odmienności celów — brak kontroli. A jest to różnica o doniosłości zasadniczej z punktu widzenia możliwości trafnej oceny skutków przedsięwzięcia. Reformator społeczny nie poświęci swego celu, a jest nim rozwiązanie problemu społecznego, dla oceny skuteczności swego programu; chyba że ocena owa będzie jego celem głównym, ale wtedy będziemy mieli już do czynienia z eksperymentem społecznym bez cudzysłowu. Tak szerokie pojmowanie eksperymentu pozwalałoby opatrywać tym mianem wszelkie reformy społeczne, podejmowane tak w skali małych społeczności, jak i społeczeństw globalnych. Jest to więc rozumienie bezużyteczne i wprowadzające tylko zamęt pojęciowy. Rzecz jasna reformy i innowacje społeczne mogą być nader wdzięcznym przedmiotem badania, w którym będzie się chciało nanieść na nie eksperymentalny schemat analizy. Nie należy jednak mieszać badania tzw. eksperymentów społecznych z badaniem problemów społecznych przy pomocy eksperymentów. Innowacje społeczne i eksperymentalne innowacje społeczne to nie to samo.

Natomiast odstąpienie od warunku drugiego, tj. czynnej manipulacji badanymi zjawiskami otworzyło drogę do nazywania eksperymentami również badania zmian naturalnych, dokonujących się bez udziału badacza a także samych tych zmian, jeśli tylko mają one wyraźne przyczyny i czytelny przebieg, co czyni •możliwą kontrolę wpływu zmian ubocznych. Zwarte migracje, wprowadzenie nowej ustawy prawnej, klęski żywiołowe, szybka industrializacja rejonu rolniczego są przykładami takich zmian. Ich badania nazywa się eksperymentami naturalnymi, przy czym dla eksperymentu *sensu stricto* rezerwuje się miano eksperymentu projektowanego lub kontrolowanego. Kiedy zmiany dyskutowanego rodzaju dokonują się na oczach badacza, mówi się o eksperymentach naturalnych *ex nunc* a gdy ich badanie ma charakter retrospektywny — o eksperymentach *ex post facto.*

Nie są to określenia fortunne. Dwuznaczna nazwa „eksperyment naturalny" sugerować może, iż eksperyment rozumiany ściśle jest sztuczny w innym sensie niż „umyślnie wykonany", co — jak zobaczymy — nie jest ani tak proste ani tak oczywiste. Natomiast określenie „eksperyment kontrolowany" podsuwa myśl,

[1] Por. K. Ajdukiewicz, Logika pragmatyczna, Warszawa 1975, s. 227 - 9; E. Nagel, Struktura nauki, Warszawa 1970, s. 387; J. K. Feibleman; Scientific method, The Hague 1972, s. 113-4.

że w badaniu zwanym eksperymentem naturalnym brak jest kontroli, co nie jest prawdą. Z punktu widzenia przyjętej w tej pracy definicji eksperymentu wyrażenie „eksperyment naturalny" jest wewnętrznie sprzeczne, zaś wyrażenia „eksperyment kontrolowany" i „eksperyment projektowany" to pleonazmy. Badania zwane eksperymentami naturalnymi nie. należą do rodziny badań eksperymentalnych a nazywanie ich tak właśnie pociąga za sobą ujemne skutki semantyczne i perswazyjne: powstaje mętlik pojęciowy a same owe badania zaczynają świecić nie swoim bynajmniej blaskiem i zyskują nieusprawiedliwione zaufanie w swą wartość dowodową. Istnieje dobry obyczaj nazywania ich pseudo- lub quasi-eksperymentami. Pozostańmy przy nim.

Tak jak rezygnacja z któregoś z warunków definiujących termin „eksperyment" prowadziła do rozszerzenia jego zakresu na zabiegi, których nie chcieliśmy nazywać eksperymentami a nawet badaniami, tak nakładanie na te warunki ograniczeń dodatkowych prowadzi do zawężania terminu „eksperyment" kosztem badań, których nie chcemy przestać nazywać badaniami eksperymentalnymi. Tę możliwość realizują ci badacze, którzy uznają badania za eksperymentalne tylko wtedy, gdy kontrola wprowadzanych zmian ma postać szczególną. Jest nią randomizacja czyli losowe przypisanie badanych jednostek poszczególnym rodzajom oddziaływań[2]. Bez wątpienia kontrola wpływu manipulacji eksperymentalnych przez randomizację jest niezmiernie efektywną i wygodną dla analizy wyników formą kontroli. Niemniej jednak jest ona tylko jedną z metod kontroli i odmawianie różnym badaniom miana eksperymentalnych tylko dlatego, że stosują inne niż randomizacja sposoby kontroli nie byłoby uzasadnione. Zbyt mocno gwałciłoby nasze obyczaje językowe i lekceważyłoby ważne podobieństwa między eksperymentami posługującymi się różnymi metodami kontroli. Warto natomiast zdawać sobie sprawę z różnic między eksperymentami z randomizacją i bez. Będzie o nich mowa dalej.

Warto na koniec zanotować jeszcze jedno, najszersze znaczenie terminu „eksperyment". Oto w przyrodoznawstwie i tych naukach o człowieku, które uprawiane są wedle jego wzorów metodologicznych, przyjęło się nazywać eksperymentowaniem całość postępowania doświadczalnego, bez względu na to, czy ma ono charakter ściśle eksperymentalny czy też wyłącznie obserwacyjny. Tak pojmowane eksperymentowanie przeciwstawiano niegdyś spekulacjom, współcześnie jest ono dopełnieniem myślenia teoretycznego i rozumowania formalnego.

Odtąd, kiedy będziemy mówić o eksperymencie, będziemy mieć na myśli tylko jego — podany wyżej — sens ścisły.

Kiedy spogląda się na eksperyment na tle innych rodzajów badań empirycznych w socjologii i gdzie indziej, to widzi się, że tym co go wyróżnia z tła jest planowy, czynny wpływ badacza na interesujące go aspekty naturalnych stanów rzeczy i ich zmian. Właściwość ta przeciwstawia tym samym badania eksperymentalne badaniom obserwacyjnym, które polegają wyłącznie na rejestracji i pomiarze zjawisk naturalnych. Takie pojmowanie obserwacji nie przesądza o tym, jakimi metodami będziemy się posługiwać w tych czynnościach. Obserwacja *sensu largo* obejmuje nie tylko obserwację w sensie ścisłym, stosowaną z powodzeniem w badaniach terenowych, lecz także analizę dokumentów urzędowych i osobistych, sondażowe badania opinii, publicznej, badania treści przekazów masowych, pomiary socjo- i psychometryczne itd., byleby tylko zjawiska badane tymi sposobami były naturalnymi stanami procesów społecznych a nie skutkami celowych manipulacji badaczy, bo i tak przecież być może.

Takie przeciwstawienie eksperymentu badaniom obserwacyjnym bywa kwestionowane z dwu stron. Z jednej strony wskazuje się, że w eksperymencie istnieją elementy obserwacji a z drugiej, że w badaniach obserwacyjnych istnieją elementy manipulacji. Przyjrzyjmy się tym argumentom dokładniej.

Istotnie, po to, aby określić następstwa manipulacji eksperymentalnych, trzeba zawsze zaobserwować stan rzeczy po manipulacji a w wielu schematach eks-

[2] Np. D. T. Campbell, J. C. Stanley, Experimental and. quasi-experimental designs for research, Chicago 1973.

perymentalnych — również przed manipulacją. Eksperyment wykonuje się po to, by móc przeprowadzić określone obserwacje.

Jednakże termin „obserwacja" użyty w tym kontekście oznacza każdą planową procedurę zbierania danych empirycznych.. Do tak rozumianej obserwacji zmierza każde postępowanie badawcze, gdyż jest ona jedynym bezpośrednim źródłem wiedzy o rzeczywistości. Jest to obserwacja *sensu* już nie *largo, lecz largissimo*. *W* eksperymencie — tak samo jak w badaniach nieeksperymentalnych — można ją wykonywać najrozmaitszymi sposobami. Można zastosować obserwację ekstraspekcyjną, wywiad, pomiar psychometryczny, spis statystyczny, można skorzystać z okresowych ocen działalności dokonywanych przez instytucje itd. Natura eksperymentu ani nie wyznacza, ani nie wyklucza żadnego z tych sposobów — czyni to natomiast problem badawczy, łatwość korzystania z poszczególnych metod czy też przyzwyczajenie eksperymentatorów. W każdym jednak razie w eksperymencie rejestruje się skutki manipulacji badacza a nie tylko naturalne stany procesu społecznego. W badaniach obserwacyjnych nieeksperymentalnych — przeciwnie.

Wiadomo też dobrze, że badacz wpływa na badane zjawiska również w wielu badaniach obserwacyjnych nieeksperymentalnych. Niekiedy wpływ ten polega na umieszczaniu badanych *w* nowej, niecodziennej sytuacji, która pozwoli na łatwiejszą obserwację. Tak postępuje psychiatra, który chcąc wydać trafną diagnozę stanu zdrowia człowieka podejrzanego o chorobę, kieruje go „na obserwację" do szpitala psychiatrycznego. Inny rodzaj wpływu wywiera badacz, który nie chcąc poprzestać na rejestracji samych tylko spontanicznych zachowań i wypowiedzi, ucieka się do interrogacji w celu wywołania nowych wypowiedzi, które będą dlań wskaźnikami interesujących go zjawisk. Ankietowe badania socjologiczne i pomiary psychologiczne oparte są na tym rodzaju wpływu. Obie wyróżnione formy oddziaływania badacza na badanych mogą zresztą występować łącznie.

Jednakże manipulacji tych nie wykonuje się po to, by ocenić ich następstwa, lecz po to, aby odsłonić ukrytą rzeczywistość bez jej naruszania. Bez wątpienia interwencje te nie pozostają bez wpływu również na te aspekty zachowania badanych, które badacza akurat interesują. Choroby jatrogenne w przypadku manipulacji pierwszego rodzaju i efekty reaktywne w przypadku manipulacji drugiego rodzaju są nazwami zmian wywołanych przez samo badanie. Niemniej jednak te metodopochodne zmiany są faktami niezamierzonymi i niepożądanymi, gdyż badacz może wziąć je za zachowania spontaniczne; mówi się wówczas, że bierze on artefakty za fakty. Zauważyć trzeba, że i owe wpływy proceduropochodne mogą stanowić przedmiot badania. Dzieje się tak w tzw. eksperymentach metodologicznych, które mają za cel opisanie ich po to. by móc je w zwykłych badaniach skutecznie eliminować.

Opozycja „eksperyment — obserwacja" zostaje więc utrzymana. Pociąga to za sobą rozróżnianie roli obserwatora i eksperymentatora. Ładnie ujął je Cuvier: „obserwator słucha przyrody; eksperymentator zapytuje ją i zmusza do odsłonięcia się".

Eksperyment i obserwacja nie są jednak sobie przeciwstawne pod względem logicznym tj. pod względem roli rozumowań i danych empirycznych oraz schematów wnioskowania w procesie sprawdzania hipotez. Logika ta wspólna jest wszystkim badaniom kontrolowanym, bez względu na ich eksperymentalny czy też czysto obserwacyjny charakter. „Badanie kontrolowane polega na rozmyślnym szukaniu dających się przeciwstawić sytuacji, w których badane zjawisko bądź występuje regularnie (w identycznej lub w różnych postaciach), bądź też występuje w pewnych przypadkach, a nie występuje w innych, oraz na badaniu pewnych czynników wyróżnionych w tych sytuacjach w celu stwierdzenia, czy zmianom tych czynników towarzyszą zmiany badanych zjawisk — przy czym zarówno owe czynniki, jak i różne formy zjawisk selekcjonuje się dla starannej obserwacji w oparciu o hipotezę, iż zachodzą między nimi istotne związki"[3]. Sytuacje porównawcze, o których mowa, możemy bądź znaleźć w życiu (obserwacja) bądź stworzyć (eksperyment)

[3] Nagel, op. cit. s. 389.

— z punktu widzenia logiki empirycznej kontroli hipotez różnica ta jest wszakże bez znaczenia; ma ona wyłącznie charakter pragmatyczny.

2. ZASTOSOWANIA I SZCZEGÓLNE ZALETY EKSPERYMENTU

Nazwanie różnic między eksperymentem a badaniami nieeksperymentalnymi różnicami pragmatycznymi w żadnym wypadku nie oznacza ich lekceważenia. Przeciwnie — to właśnie pragmatyczne aspekty eksperymentu decydują o jego szczególnej wartości dla badań empirycznych. Rozważymy momenty składające się na tę wartość.

2.1. EKSPERYMENT METODĄ SPRAWDZANIA HIPOTEZ PRZYCZYNOWYCH

Siła eksperymentu ujawnia się przede wszystkim w sprawdzaniu hipotez przyczynowych — eksperyment jest jego doskonałym narzędziem.

2.1.1. Pojęcie hipotezy przyczynowej i związku przyczynowego

Hipotezy przyczynowe zajmują wybitne miejsce wśród hipotez sprawdzanych w naukach empirycznych i takież miejsce wśród twierdzeń przyjętych przez te nauki jako sprawdzone zajmują prawa przyczynowe. Jedne i drugie orzekają, tyle że z różnym stopniem asercji, o związkach przyczynowych między zdarzeniami (zmiennymi). Ale czym jest związek przyczynowy?[4]

Pojęcie związku przyczynowego jest jednym z bardziej niejasnych w słowniku nauki. Niegasnącą żywotność wykazują też spory o status zasady przyczynowości w nauce. Powiemy, że między dwoma zdarzeniami *A* i *B* zachodzi związek przyczynowy, gdy:

a) Jedno ze zdarzeń, *A,* jest warunkiem wystarczającym bądź warunkiem niezbędnym, bądź warunkiem wystarczającym i niezbędnym zarazem, bądź wreszcie niezbędnym składnikiem warunku wystarczającego do zajścia drugiego ze zdarzeń. B.

W naukach empirycznych a w naukach społecznych w szczególności rzadko lub wręcz nigdy nie formułuje się twierdzeń podających warunki wystarczające lub wystarczające i konieczne zarazem. Zdarzenie wyróżniane jako przyczyna jest zwykle koniecznym składnikiem warunku wystarczającego; mówi się też, że jest ono warunkiem dopełniającym do warunku wystarczającego.

Tak rozumiane zdarzenie *A* dopełnia pewien warunek C do warunku wystarczającego do zajścia B; *A* i C tworzą rarazem warunek wystarczający do wystąpienia B. Ani samo C ani samo *A* nie wystarczają do zajścia B. Możliwe są tu różne sytuacje. Warunek *A* może być składnikiem niezbędnym wielu warunków wystarczających tzn. może dopełniać do warunku wystarczającego różne warunki C*1*, *C2, ..., Cn.* Mogą istnieć alternatywne warunki dopełniające dany warunek C do warunku wystarczającego a może być i tak, że *A* jest jedynym warunkiem dopełniającym.

b) Między zdarzeniami zachodzi pewien porządek czasowy, taki mianowicie, że zdarzenie wyróżnione jako przyczyna występuje wcześniej niż zdarzenie wyróż-

[4] Odpowiedź na to pytanie opieramy na nast. pracach: J. Karpiński, Wstęp do metodologii nauk społecznych, mat. powiel. (Politechnika Wrocł. 1976), H.M. Blalock, Jr., Causal inferences in nonexperimental research, Chapel Hill 1964, rozdz. 1; E. Nagel, Struktura nauki, Warszawa 1970, s. 73 - 4 , 479, 498 - - 503; K. Ajdukiewicz, Logika pragmatyczna, Warszawa 1975, s. 155-6; S. Nowak, Understanding and prediction, Dordrecht 1976, rozdz. 5.

nione jako skutek. Aby porządek ten można było w ogóle ustalić, zdarzenia muszą być czasowo określone.

Są to minimalne warunki definicyjne związku przyczynowego. Sprowadzają one go do stałego następstwa zdarzeń. Wielu metodologów a zwłaszcza logików uważa, że wystarczą one do zdefiniowania związku przyczynowego.

c) Wielu innych metodologów, zwłaszcza tych bliskich badaniom empirycznym, uważa jednak, iż stałe następstwo nie wystarcza do zdefiniowania związku przyczynowego, aczkolwiek jest do zdefiniowania konieczne. Ich zdaniem, warunkiem związku przyczynowego jest *wywoływanie* zdarzenia wyróżnionego jako skutek przez zdarzenie wyróżnione jako przyczyna. Takie jest i nasze stanowisko.

Związek przyczynowy jest nie dowolnym stałym następstwem zdarzeń, lecz tylko takim następstwem, w którym jedno zdarzenie wywołuje, produkuje czy generuje drugie. Ów dynamiczny związek między zdarzeniami nie da się jednak wyrazić w tym języku, w którym zostały sformułowane dwa poprzednie warunki. W języku tym faktycznie nie można powiedzieć o związku przyczynowym więcej niż powiedziano. Tymczasem pojęcie związku przyczynowego nie jest pojęciem logicznym, lecz teoretycznym; ontologicznym — jak powiadają niektórzy autorowie. Związek ten nie jest bezpośrednio obserwowalny. W badaniach empirycznych w gruncie rzeczy nic można stwierdzić nic więcej, jak tylko stałe następstwo. O istnieniu związku przyczynowego wnioskuje się ze stałości następstwa. Do tego jednak, by uzasadnić taką inferencję potrzebne są dodatkowe założenia o innych zmiennych, takie jak założenie o izolacji badanych zmiennych czy klauzula *ceteris paribus*. Te założenia wzięte są skądinąd.

Zdefiniowane przez odwołanie się do związku przyczynowego twierdzenia przyczynowe mogą posiadać różny stopień . ogólności teoretycznej i kulturowo-historycznej. Będąc warunkowymi twierdzeniami ogólnymi o związkach między zdarzeniami, mogą pozostać tylko uogólnieniami przypadkowymi a mogą też zasłużyć sobie na godność praw teoretycznych. Twierdzenia przyczynowe o danym stopniu ogólności teoretycznej mogą odnosić się do pewnych tylko, wymienionych z nazwy kultur, regionów i okresów i wówczas będą generalizacjami historycznymi. Jeśli jednak uda się określić ich zasięg bez uciekania się do imion własnych, staną się one uogólnieniami (prawami) uniwersalnymi. Oba rodzaje ogólności — teoretyczna i historyczno-kulturowa idą zazwyczaj ze sobą w parze.

Potrzeba posługiwania się przez naukę terminem „przyczyna" i pochodnymi była w filozofii nauki namiętnie dyskutowana. Spośród poglądów — nieraz krańcowych — wygłaszanych w tej kwestii najbliższe nam jest rozsądne i wyważone stanowisko Ernestą Nagla. Przyznając rację tym, którzy twierdzą, że termin „przyczyna" wyjątkowo tylko spotyka się w rozwiniętych naukach, Nagel pisze, iż „aczkolwiek termin ten może być nieobecny, to idea, do której on się odnosi jest wciąż w szerokim obiegu. Obfituje w nią nie tylko mowa potoczna, lecz także badania spraw ludzkich prowadzone przez ekonomistów, psychologów społecznych i historyków. Szerzy się ona również w dostarczonych przez przyrodoznawców sprawozdaniach z ich zabiegów laboratoryjnych, tak samo jak i w interpretacjach formalizmu matematycznego, proponowanych przez fizyków-teoretyków. Pojęcie przyczynowości, choć niekoniecznie słowo «przyczyna« jest także obecne wtedy, gdy naukowcy odróżniają w rozmaitych badaniach korelacje pozorne i rzeczywiste. Krótko, idea przyczyny nie jest w nowoczesnej nauce tak przestarzała, jak niekiedy się sądzi"[5].

Tyle o pojęciu związku i prawa przyczynowego.

2.1.2. Ogólny schemat sprawdzania hipotez przyczynowych

Z przedstawionej charakterystyki związku przyczynowego wynikają pewne konsekwencje operacyjne dla badań nakierowanych na sprawdzanie hipotez przyczynowych. Do tego mianowicie by przyjąć hipotezę, że zdarzenie (zmienna) *A* jest przyczyną zdarzenia (zmiennej) *B* konieczne są trzy operacje:

[5] E. Nagel, Types of causal explanation in science, w: D. Lerner (ed.), Cause and effect, New York 1965, s. 12.

a. Ustalenie współwystępowania i współzmienności *A* i *B* Jeśli potencjalna przyczyna i potencjalny skutek nie występują i nie zmieniają się razem, to żadne' z nich nie może być przyczyną drugiego.

b. Wykazanie, że *A* wyprzedza *B* a w każdym razie nie jest odeń późniejsze. Jeśli potencjalna przyczyna występuje po potencjalnym skutku, to nie może ona być przyczyną faktyczną.

c. Dowiedzenie, że żadne inne zdarzenia Współwystępujące z *A* nie mogły wywołać B. Dopóki nie odrzuci się możliwości wyjaśnienia B innymi niż *A,* lecz współwystępującymi z *A* i B zdarzeniami, dowód przyczynowy nie będzie pełny a więc i pewny.

Ustalenie współwystępowanie dwóch zmiennych jest rzeczą stosunkowo łatwą — wystarczą do tego badania korelacyjne. Wykażą one np. dodatnią korelację rygoryzmu obyczajowego z wiekiem a ujemną z wykształceniem badanych, korelację między kontaktem ze źródłem perswazji a akceptacją propagowanych treści czy zbieżność między stopniem przyjaźni w stosunkach między ludźmi a podobieństwem ich poglądów. Jednakże zbieżność nie dowodzi jeszcze zależności (przyczynowej). Ustalone korelacje mogą być wynikiem działania trzecich zmiennych — nazywa się je wówczas pozornymi, choć — jak trafnie zauważono[6] — nie ma pozornych zależności, są tylko pozorne interpretacje. Obserwowana korelacja może brać się stąd, że obie zmienne są skutkami tej samej przyczyny. Zarówno kontakt z pewnym źródłem perswazji, jak i uznawanie upowszechnianych przez nie treści mogą zależeć od wykształcenia badanych, same nie będąc przyczynowo związanymi ze sobą. Ustalona korelacja może też wynikać z tego, że domniemana zmienna niezależna jest skorelowana z wieloma innymi zmiennymi i któraś z tych zmiennych a nie ona, jest faktyczną przyczyną drugiej ze zmiennych. Taki charakter może mieć związek wieku i rygoryzmu obyczajowego. Wiek, skorelowany jest z wykształceniem a wykształcenie ma wpływ na stopień rygoryzmu. Były to dwie najprostsze sytuacje. Prawda, że wiele ze stwierdzonych korelacji, jak choćby wykształcenia z tolerancją, ma charakter rzeczywisty a nie pozorny. Jednakże samo ustalenie współwystępowania między zmiennymi nie wystarcza do uznania jednej ze zmiennych za przyczynę drugiej. Konieczne są operacje dodatkowe.

Operacje te polegają na wysuwaniu i sprawdzaniu hipotez wyjaśniających zmienną zależną B innymi zmiennymi niż zmienna *A,* figurująca jako zmienna niezależna w hipotezie początkowej. Hipotezy te są więc konkurencyjnymi względem hipotezy sprawdzanej. Wytłumaczenie uznawania pewnych poglądów wykształceniem a nie oddziaływaniem perswazyjnym oraz wytłumaczenie rygoryzmu niskim wykształceniem a nie wiekiem były przykładami takich hipotez. Dopiero wtedy, gdy odrzuci się hipotezy konkurencyjne, można przyjąć hipotezę początkową. Często tym, co przyjmuje się, jest jedna z hipotez konkurencyjnych, a tym co się odrzuca, jest hipoteza początkowa. Przedstawimy te procedury dokładniej.

Truizmem jest twierdzenie, że hipotezy przyczynowej nie można, definitywnie potwierdzić — definitywnie można ją tylko obalić z całą pewnością można dowieść tylko braku związku przyczynowego między zmiennymi. Istnienia tego związku nigdy nie można dowieść w sposób ostateczny, można je tylko uczynić mniej lub bardziej prawdopodobnym. Dzieje się tak dlatego, że logiczne schematy wnioskowania z faktów zgodnych z hipotezą o jej prawdziwości są schematami zawodnymi, natomiast schematy wnioskowania z faktów niezgodnych z hipotezą o jej fałszywości są schematami niezawodnymi.

Nasuwają się zatem dwa sposoby podnoszenia wiarygodności sprawdzanej hipotezy[7].

Pierwszy polega na szukaniu takich faktów zgodnych z hipotezą, których wartość uzasadniająca jest szczególnie wysoka. Przy braku świadectw negatywnych

[6] M. Rosenberg, The logic of survey analysis, New York 1968, s. 28.

[7] Zob. Ajdukiewicz, op. cit., s. 367—83; C. G. Hempel, Podstawy nauk przyrodniczych, Warszawa 1968, rozdz. 4.

stopień potwierdzenia hipotezy jest tym wyższy, im więcej faktów okazało się zgodnymi z hipotezą, im bardziej były one urozmaicone, im dokładniej wyniki ich obserwacji odpowiadały przewidywaniom wysnutym z hipotezy; szczególną wartość posiadają przy tym takie fakty zgodne z hipotezą, które nie były znane w chwili jej formułowania oraz takie, których zajście wydawało się na gruncie wiedzy dotychczasowej, nie obejmującej hipotezy, mało prawdopodobne.

Inny sposób podwyższania wiarygodności sprawdzanej hipotezy polega na obalaniu hipotez konkurencyjnych. Obalenie hipotez konkurencyjnych względem hipotezy sprawdzanej wraz z niemożliwością obalenia jej samej czyni ją bardziej wartościową niż konkurentki. Stopień wiarygodności hipotezy jest tym wyższy, im więcej poważnych tj. początkowo wysoce prawdopodobnych hipotez konkurencyjnych zostało odrzuconych. Zatem, chcąc przyjąć pewną hipotezę, trzeba starać się obalić w rzetelny sposób hipotezy konkurencyjne. Jeśli próby te skończą się niepowodzeniem, hipotezę sprawdzaną możemy przyjąć jako dobrze — na danym poziomie rozwoju nauki — uzasadnioną.

Eksperymentalne sprawdzanie hipotez przyczynowych opiera się na drugiej z naszkicowanych metod.

Logicznych podstaw dla uzasadnienia hipotez przyczynowych przez eliminacje hipotez konkurencyjnych dostarcza logika indukcji eliminacyjnej. Reguły wnioskowania znane jako kanony Milla są niczym innym jak schematami eliminacji hipotez przyczynowych. Oto ich krótki uproszczony wykład, który uwypukli taki właśnie ich charakter[8],

Pytanie, na które odpowiedzieć ma wnioskowanie wedle kanonów Milla brzmi: Które z wyróżnionych zdarzeń towarzyszących zajściu zdarzenia Y jest jego przyczyną; *A, B. C* czy D? Założenie tego pytania głosi, że poszukiwana przyczyna znajduje się wśród zdarzeń wyodrębnionych i wyznacza jednoznacznie zajście Y. Problem wyjściowy można też sformułować w języku hipotez, będących próbnymi odpowiedziami na pytania o przyczynowe znaczenie poszczególnych zdarzeń. Dysponujemy zespołem konkurencyjnych względem siebie hipotez wyjaśniających zajście zdarzenia Y, przy czym hipoteza dodatkowa głosi, że jedna z tych hipotez jest prawdziwa. Pytanie, która.

Przypuśćmy, że sprawdzamy hipotezę, że zdarzenie *A* jest przyczyną Y; przypuszczamy, iż jest prawdziwa. Wiemy, że samo współwystępowanie zdarzeń Y i *A* niczego jeszcze nie dowodzi, gdyż także inne, konkurencyjne zdarzenia towarzyszą zachodzeniu Y. Aby móc »hipotezę sprawdzaną przyjąć, musielibyśmy hipotezy konkurencyjne odrzucić jako nietrafne. Kanony Milla są schematami takich eliminacji, wyznaczonymi przez definicje przyczyny.

Jeśli przyczynę zdefiniować jako warunek wystarczający, to wynika stąd, że *B* nie będzie przyczyną Y, jeśli zdarzy się, że *13* wystąpi a nie wystąpi Y. Chcąc więc wykazać, że *B* nie jest przyczyną Y, trzeba taką sytuację znaleźć a wiadomo będzie z całą pewnością, że to nie B jest warunkiem wystarczającym Y. Jeśli natomiast przyczynę rozumieć jako warunek konieczny, to B nie będzie przyczyną Y, jeśli zdarzy się, że zajdzie Y mimo braku B. Chcąc zatem wykazać, że B nie jest przyczyną Y, trzeba znaleźć taki układ zdarzeń a będzie można wyciągnąć niezbity wniosek, że to nie B jest warunkiem koniecznym Y.

Tym sposobem wyeliminowaliśmy jedną z hipotez konkurencyjnych względem hipotezy sprawdzanej i tym samym podnieśliśmy stopień potwierdzenia hipotezy sprawdzanej. W ten sam sposób postępuje się z pozostałymi hipotezami konkurencyjnymi. Ich eliminację można prowadzić krok po kroku aż do ewentualnego wyeliminowania wszystkich. Można ją też przeprowadzić od razu, jeśli tylko dysponuje się układami zdarzeń, zakładanymi przez kanon jedynej różnicy i jedynej zgodności.

Kanon jedynej różnicy wymaga takich dwu serii zdarzeń, w których jedna polega na zajściu zdarzenia Y oraz wszystkich zdarzeń podejrzanych o jego wywołanie

[8] Opiera się on na nast, pracach: M. R. Cohen, E. Nagel, An introduction to logic and scientific method, New York 1934, rozdz. 13; K. Szaniawski, O indukcji eliminacyjnej, w: Fragmenty filozoficzne, Seria II, Warszawa 1959; Ajdukiewicz: op. cit., s. 150 - 70.

a druga polega na braku zdarzenia Y i zajściu wszystkich zdarzeń podejrzanych o wpływ oprócz jednego zdarzenia A.

Żadne ze zdarzeń, które zaszły zarówno wtedy, gdy zaszło Y, jak i wtedy, gdy Y nie wystąpiło, nie może być na mocy definicji przyczyną zdarzenia Y rozumianą jako warunek wy

$$\begin{array}{cccc} A & B & C & Y \\ A & B & C & \overline{Y} \end{array}$$

starczający. Przyczyną tą może być tylko *A*. Odrzuciliśmy w ten sposób wszystkie hipotezy konkurencyjne względem hipotezy sprawdzanej, czyniącej *A* przyczyną Y. Czy znaczy to, że *A* jest przyczyną Y, czyli że sprawdzana hipoteza jest trafna? Odpowiedź twierdząca na to pytanie wynika logicznie ze wstępnej przesłanki, że szukana przyczyna faktycznie znajduje się wśród zdarzeń wyodrębnionych. Czy odpowiedź ta jest prawdziwa, jest rzeczą trafności owej przesłanki.

Od przesłanki tej wychodzi również wnioskowanie wedle kanonu jedynej zgodności. Nakazuje on szukać takich kombinacji zdarzeń, w których zajściu Y będzie stale towarzyszyć tylko jedno zdarzenie *A*, wszystkie inne zaś wystąpią tylko w niektórych przypadkach.

$$\begin{array}{cccc} A & B & C & Y \\ A & B & C & Y \\ A & B & \overline{C} & Y \\ A & \overline{B} & \overline{C} & Y \end{array}$$

Żadne ze zdarzeń, które nie występowały zawsze gdy zachodziło Y, nie może być na mocy definicji przyczyną zdarzenia Y rozumianą jako warunek konieczny. Przyczyną tą może być tylko *A*. Obaliliśmy tym sposobem wszystkie hipotezy konkurencyjne wobec hipotezy sprawdzanej głoszącej, że *A* jest przyczyną Y. Analogicznie jak w wypadku kanonu jedynej różnicy, przesłanka początkowa pozwalała przyjąć hipotezę sprawdzaną jako uzasadnioną. Prawomocność tej decyzji zależy jednak od trafności tej przesłanki. W każdym razie hipotezę możemy przyjąć jako nieobaloną i lepszą od konkurentek.

Takie ujęcie kanonów Milla uwypukla ich charakter eliminacyjny. Są one regułami sprawdzania hipotez przyczynowych przez eliminację hipotez konkurencyjnych. Podkreśla też ich charakter dedukcyjny. Są one takimi schematami rozumowań, w których wnioski wynikają logicznie z przesłanek. Jest to prawdziwe tak dla wniosków oznaczających eliminację hipotez, jak i ich akceptację. Jednak w pierwszym wypadku prawdziwość wniosku nie zależy od wiedzy przedmiotowej, co najwyżej od poprawności obserwacji. W drugim natomiast prawdziwość wniosku wyznaczona jest przez prawdziwość przyjętych przesłanek faktycznych. Utrzymują one, że poszukiwane przyczyny pewnych zdarzeń znajdują się wśród zdarzeń wyodrębnionych i że, co więcej, są one przyczynami jedynymi. Obecność założeń pozalogicznych, teoretycznych jest trzecią cechą przedstawionego wykładu kanonów Milla. Wiedza przedmiotowa potrzebna jest także do tego, by stosowanie tych kanonów było przedsięwzięciem w ogóle wykonalnym praktycznie. Nie byłoby takim, gdyby w skład zespołu zdarzeń uważanych za możliwe przyczyny wchodziło zbyt wiele zdarzeń faktycznie nieistotnych. Wnioskowania wedle kanonów indukcji eliminacyjnej muszą więc być poprzedzone poważną robotą teoretyczną.

W żadnym razie kanony te nie są „maszyną do wykrywania praw przyrody, którą mógłby obsługiwać pierwszy lepszy”[9].

Wykazanie przyczynowego charakteru związku między zdarzeniami jest niewątpliwie ważnym osiągnięciem, nie pozwala jednak jeszcze wskazać jednego z dwóch zdarzeń jako przyczyny, a drugiego jako jej skutku. Do tego niezbędne jest uprzednie ustalenie porządku czasowego między zdarzeniami. Innymi słowy, wymagane jest odrzucenie jeszcze jednej hipotezy konkurencyjnej, sugerującej odwrotny kierunek wpływu przyczynowego, niż to zakłada hipoteza sprawdzana. Same kanony Milla nic tu nie pomogą, bo stosują się równie dobrze do wykrywania przyczyn jak i do ustalania skutków. Określenie następstwa czasowego między powiązanymi przyczynowo zmiennymi jest w badaniach społecznych często sprawą oczywistą, zwłaszcza w wypadkach zależności między cechami społeczno-demograficznymi a postawami i zachowaniami. Ale bywa i poważnym problemem. Nawet gdyby dowieść przyczynowego charakteru związku między kontaktem ze źródłem perswazji a posiadaniem propagowanych przez nie poglądów, to i tak nie byłoby wiadomo, skąd ten związek sie bierze, czy z wpływu perswazji na opinie czy z poszukiwania kontaktu z treściami zgodnymi z posiadanymi już poglądami. Równie trudno odpowiedzieć *a priori,* czy zaobserwowana zbieżność poglądów u przyjaciół jest wynikiem upodabniania się przyjaciół czy też dobierania się przyjaciół według podobieństwa poglądów.

Dopiero wtedy, gdy stwierdzimy, że hipotetyczna przyczyna pewnego zdarzenia współwystępuje z nim, gdy wykażemy, że współwystępowanie to nie ma charakteru pozornego i gdy ustalimy, że hipotetyczna przyczyna występuje wcześniej niż to zdarzenie, będziemy mieli prawo przyjąć sprawdzaną hipotezę przyczynową.

2.1.3. Eksperymentalne sprawdzenie hipotez przyczynowych

Najlepszym sposobem wykonania zadań koniecznych dla przyjęcia hipotezy przyczynowej byłoby badanie takich kombinacji zdarzeń, które spełniają warunki kanonów indukcji eliminacyjnej i — co więcej — tworzą się na oczach badacza. W zestawieniach takich jedyną różnicą czy zgodnością byłyby zdarzenia wskazane w poprzednikach sprawdzanych hipotez, obserwacja ustalałaby zaś czy wystąpiły także zdarzenia objęte ich następnikami. Stwierdzone w ten sposób zależności byłyby przyczynowymi z założenia — możliwość stwierdzenia zależności pozornych byłaby wykluczona przez charakter układu dobrze kontrolowanego. Wreszcie, obserwacja układu zdarzeń w trakcie jego formowania się, pozwalałaby łatwo określić, które ze zdarzeń jako pierwsze w czasie jest przyczyną a które, jako późniejsze jest skutkiem. Następstwo czasowe zdarzeń byłoby obserwowane bezpośrednio.

Jednakże spontaniczne, naturalne powstawanie takich układów „jest to usługa, jaką natura rzadko jest tak uprzejma nam wyświadczać, jako że zbudowana jest na schemacie zupełnie różnym od tego, jaki ułatwia nam nasze badania” — zauważa smętnie sam Mill[10]. Układy naturalne są skomplikowane w tym sensie, że są od siebie różne lub do siebie podobne pod zbyt wieloma względami, by móc w każdym wypadku ustalić przyczynowe znaczenie poszczególnych, pojedynczych zdarzeń. Zdarzenia te mają tendencję do występowania we względnie trwałych związkach z innymi zdarzeniami, łatwo więc wziąć skutki jednych zdarzeń za skutki zdarzeń im towarzyszących. Aby pokonać te trudności, trzeba uciec się do utworzenia wygodnych dla indukcji sytuacji w drodze „sztucznego zestawienia okoliczności” czyli eksperymentowania.

Na eksperymentowanie składa się stymulacja, kontrola i obserwacja. Stymulacja polega na wprowadzaniu zmian, których następstwa chce się ocenić, czyli bodźców. Kontrola ma umożliwić trafne przypisanie następstw zmianom czyli reakcji bodźcom. Obserwacja ma następstwa te zarejestrować. Brak któregokolwiek z tych elementów czyniłby eksperyment niemożliwy lub bezsensowny, ale kluczową rolę w sprawdzaniu hipotez przyczynowych pełni kontrola.

---

[9] K. Ajdukiewicz, Franciszek Bacon z Werulamu. Dzieło i życie. „Wstęp” do: F. Bacon, Novum Organum, Warszawa 1955. s. XLII; uwaga ta dotyczy tablic Bacona, pierwowzoru kanonów Milla.

[10] J.S. Mill, System logiki, Warszawa 1962, t. 1 s. 591.

Kontrola polega na tworzeniu układów porównawczych w stosunku do tych układów, które poddaje się działaniu bodźców lub — rzadziej — na izolacji bodźców od ich możliwych zakłóceń. Układ porównawczy może być pozostawiony sobie, nie poddawany stymulacji; nazywa się go wtedy grupą kontrolną. Jest to najprostsza forma kontroli. Eksperyment ma jednak bardzo często na celu porównanie wpływów różnych rodzajów bodźców. Wtedy układy poddane poszczególnym rodzajom bodźców stanowią dla siebie nawzajem układy porównawcze; dodatkowy układ może być nie poddawany żadnym oddziaływaniom.

Kontrola wymaga ujednolicenia wszystkich grup ze względu na wartości zmiennej zależnej i zmiennych istotnych, tj. takich, które mogłyby wpływać na zmienną zależną. Kontrola wymaga również tego, by podczas eksperymentu oddziaływania inne niż bodziec były jednakowe we wszystkich grupach.

Jeśli po wprowadzeniu bodźca zaobserwuje się różnice w wartościach zmiennej zależnej w poszczególnych grupach, to:

1. nie będzie można wytłumaczyć ich różnicami w wartościach zmiennej zależnej przed działaniem bodźca, bo takich różnic nie było;

2. nie będzie można wyjaśnić ich różnicami w wartościach trzecich zmiennych, przed działaniem bodźca, bo (takich różnic też nie było; zmienne te nie były skorelowane z bodźcami, nie mogą więc tłumaczyć reakcji;

3. nie będzie można wytłumaczyć ich różnicami innych niż bodziec oddziaływań, bo i takich różnic nie było.

W ten sposób, wnioskując wedle kanonu jedynej różnicy, eliminuje się wiele możliwych hipotez konkurencyjnych względem hipotezy sprawdzanej. Żadne ze zdarzeń, które zachodziły zarówno wtedy, gdy zaobserwowaliśmy reakcję, jak i wtedy, gdy jej nie stwierdziliśmy, nie może być przyczyną tej reakcji. Może nią być tylko różnica a tylko zmienna eksperymentalna przybierała różne wartości. Ujednolicenie grup eksperymentalnych i kontrolnych czyni bardzo wysoce prawdopodobną tezę, że bodziec był różnicą jedyną. Pewność ta zależy od tego, czy się wyrównało grupy pod wszystkimi istotnymi względami i czy zapewniło się faktycznie jednolitość oddziaływań pozastymulacyjnych. Widać więc, że w eksperymencie łatwo stosować zasady indukcji rozumianej nie tylko jako wnioskowanie eliminujące, ale i jako wnioskowanie uzasadniające.

W badaniach eksperymentalnych łatwo nie tylko unikać brania zależności pozornych za rzeczywiste. Łatwo również ustalać porządek czasowy między powiązanymi przyczynowo zdarzeniami. Następstwo to obserwowane jest bezpośrednio. To, co jest przyczyną określa sam eksperymentator — jest nim jego manipulacja. W ten sposób eliminuje się dodatkową hipotezę konkurencyjną względem sprawdzanej — głosiłaby ona, że ustalony związek przyczynowy jest wynikiem. wpływu zmiennej, określonej w hipotezie sprawdzanej jako zmienna zależna a nie jako zmienna niezależna.

Eksperyment kontroluje skutecznie nie tylko zmienne uboczne wobec zmiennej niezależnej, ale także jej składowe. Wiele ważnych zmiennych w badaniach społecznych ma charakter złożony. Religijność, osobowość autorytarna, przynależność klasowa są przykładami takich zmiennych. Stwierdzenie, że któraś z tych zmiennych warunkuje pewne zachowania czy postawy, rodzi więc pytanie, które składowe tej zmiennej mają wpływ na zmienną zależną. Ten, kto ustalił, że inteligenci częściej niż robotnicy cenią bardziej to czy owo, staje wobec pytania: która z różnic między tymi grupami odpowiedzialna jest za różnice preferencji. Wykształcenie, rodzaj pracy, prestiż, dochód nasuwają się jako konkurencyjne odpowiedzi. Przypuśćmy, że okazało się, iż istotną składową jest wykształcenie. Ale tu pojawia się kolejne pytanie: która ze składowych wykształcenia — zasób wiedzy czy umiejętność samodzielnego myślenia? Są to poważne problemy badawcze, trudne często do rozwiązania z powodu niemożliwości dowolnego oddzielania w warunkach naturalnych zespolonych składowych zmiennych złożonych.

Badania eksperymentalne pozwalają natomiast na systematyczne specyfikowanie złożonych zmiennych niezależnych i określenie przyczynowej doniosłości poszczególnych ich aspektów. Cel ten można osiągnąć albo przez wyizolowanie interesujących aspektów zmiennej niezależnej, albo przez tworzenie nowych grup kontrolnych w celu wyrównania jednych aspektów tak, by można było precyzyjnie określić wpływ drugich. Logika kontroli jest tu więc taka sama, jak przy kontroli zmiennych ubocznych.

Eksperyment jest badaniem powtarzalnym i faktycznie powtarzany bywa częściej niż inne badania. Powtórzenie eksperymentu stanowi dodatkowy sprawdzian trafności jego wyników.

Na wystąpienie lub wielkość zaobserwowanej w eksperymencie zmiany wpływ mają nie tylko te aspekty zastosowanych bodźców, które stanowiąc operacyjne odpowiedniki pewnych poziomów zmiennej niezależnej, figurującej w sprawdzanej hipotezie, są teoretycznie istotnymi aspektami manipulacji. Już w klasycznych badaniach w Hawthorne grupa Eltona Mayo zauważyła, że wpływ eksperymentalnej poprawy oświetlenia na wydajność pracy nie da się zredukować do wpływu wzrostu siły światła, gdyż sam fakt zmiany i wzrost zainteresowania robotnikami przyczynia się do wzrostu wydajności pracy. W eksperymencie tym tylko wzrost siły światła był teoretycznie istotnym aspektem manipulacji, natomiast sam fakt zmiany i wzrost zainteresowania robotnikami były aspektami teoretycznie nieistotnymi[11].

Obok teoretycznie nieistotnych aspektów manipulacji, w eksperymencie obecne są także inne teoretyczne nieistotne czynniki. Obok konkretnego sposobu przekładu występujących w hipotezie zmiennych niezależnych na manipulacje, takimi czynnikami są: warunki, w których eksperyment się odbywa, sposób jego aranżacji, szczególne cechy populacji, w której się go prowadzi, osoba eksperymentatora, metoda pomiaru zmiennej zależnej. Nie wynikają one jednoznacznie ze sprawdzanej hipotezy, lecz z rozmaitych okoliczności sytuacyjnych, „przypadkowych". Zmienne niezależne można operacjonalizować na wiele sposobów, eksperyment można prowadzić w różnych warunkach, można go aranżować w różny sposób itd. Każdy z tych czynników może jednak posłużyć do konkurencyjnego wyjaśnienia stwierdzonej zmiany. Może zachowanie badanych nie jest reakcją na teoretycznie istotne aspekty bodźca, lecz na jego aspekty nieistotne, na cechy otoczenia, eksperymentatora itd.? W badaniach w Hawthorne okazało się np., że pogorszenie oświetlenia również powoduje wzrost wydajności pracy, z czego wynika, że wzrost ten zależy właśnie od teoretycznie nieistotnych aspektów manipulacji. Możliwości takich konkurencyjnych wyjaśnień zmiany eksperymentalnej muszą więc być wyeliminowane.

Od omawianych zanieczyszczeń bodźca uwolnić się nie sposób, gdyż nie ma ani doskonałych operacjonalizacji ani neutralnych warunków, ani idealnie trafnych metod pomiaru. Można je tylko zmniejszać. Nade wszystko jednakże trzeba je maksymalnie różnicować. Jest to cel powtórzenia. Nie chodzi tu o dokładne powtarzanie eksperymentu — byłoby to całkowicie bezsensowne. Jeśli w różnych warunkach, przy różnych sposobach przekładu teoretycznej zmiennej niezależnej na manipulacje, przy różnych zbiorowościach, różni eksperymentatorzy, mierząc zmienną zależną różnymi metodami uzyskają te same wyniki, to znaczy, że żaden z tych czynników nie może wytłumaczyć zaobserwowanych zmian. Może to czynić tylko zmienna eksperymentalna. Pozytywna część tego wniosku jest tym pewniejsza, im bardziej wszechstronne było powtórzenie.

Gdy z identyczności uzyskanych podczas powtórzenia wyników wnosimy o nieistotności różnicowanych czynników a o istotności zmiennej eksperymentalnej, to opieramy się na kanonie jedynej zgodności.

Nowoczesne techniki eksperymentalne pozwalają rozwiązywać bardziej złożone problemy przyczynowe niż to czyniły schematy oparte na kanonach Milla.

[11] Zob. J.A.C. Brown, Społeczna psychologia przemysłu, Warszawa 1962, rozdz. 3; dla ścisłości, eksperyment ten nie był jeszcze dziełem Mayo — była nim tylko jego interpretacja.

Pozwalają one badać przyczynowy wpływ ’ wielu zmiennych równocześnie i interakcje — warunkowanie wpływu jednych zmiennych przez drugie — uwzględniają tak zmienne jakościowe jak i ilościowe i liczą się z czynnikami niekontrolowanymi czyli błędem eksperymentalnym. Oparte są one na randomizacji — losowym przydziale obiektów badanych do poszczególnych odmian bodźców i na analizie wariancji uzyskanych wyników. Analiza wariancji polega na podziale całkowitej wariancji wyników indywidualnych na wariancje składowe, stanowiące miary zmienności powodowanej przez poszczególne źródła zmienności: pojedyncze czynniki, ich kombinacje, czynniki niekontrolowane. Ocenia się w ten sposób wpływ przyczynowy tych źródeł zmienności. Dokładniej będzie mowa o tym w dalszej części pracy.

Zamykając rozważania nad szczególną wartością eksperymentu dla sprawdzania zależności przyczynowych, zauważmy, że miarą tej wartości jest skłonność do operacyjnego definiowania związku przyczynowego jako takiego związku, którego nie wyeliminował dobrze kontrolowany eksperyment. Definicję tę wolno traktować tylko jako częściową, zrelatywizowaną do badań eksperymentalnych.

2.2. INNE ZASTOSOWANIA EKSPERYMENTU

Funkcje eksperymentu nie sprowadzają się do sprawdzania hipotez przyczynowych. Aczkolwiek bezpośrednim celem eksperymentu jest zawsze obserwacja następstw manipulacji eksperymentatora, to cel ten może być podporządkowany różnym celom nadrzędnym. Sprawdzanie hipotez przyczynowych było jednym z nich, niewątpliwie najczęściej realizowanym. Bywają i cele inne.

Eksperyment służy odkrywaniu nowych zjawisk i stawianiu nowych hipotez. Jeśli jest to jego celem nadrzędnym, to nazywa się go heurystycznym. Jeśli jest to jego produktem ubocznym, to mówi się, że pełni on funkcje heurystyczne obok funkcji testowych.

Czasami zmiany wprowadza się nie po to, by sprawdzić, czy ich następstwa będą zgodne z hipotezą, lecz po to, by zobaczyć, jakie te następstwa będą. Raz jest się kierowanym pytaniem rozstrzygnięcia, drugi raz pytaniem dopełnienia: „co z tego wyniknie?” Nie zawsze bowiem znajomość badanej dziedziny pozwala na sformułowanie hipotez. Wtedy można i trzeba wykonać eksperyment heurystyczny.

W eksperymentach często stwierdza się zjawiska nie przewidziane ani przez hipotezę, ani przez eksperymentatora. Gdyby ten mógł wszystko przewidzieć, to zamiast eksperymentować, powinien raczej kalkulować. Rad, penicylina, zdanie sobie sprawy z wpływu nieformalnych stosunków w zakładzie pracy na jego funkcjonowanie są klasycznymi przykładami odkryć, będących ubocznym produktem badań.

Eksperyment zwykle nie tylko odpowiada na pytania postawione, lecz zmusza do postawienia nowych. Niektórzy metodologowie mówią nawet, że dobry eksperyment winien więcej pytań stawiać niż rozstrzygać. W badaniach społecznych wyjątkowo tylko zdarza się, że wszyscy badani reagują na bodziec w jednakowy sposób, a zatem natychmiast powstaje pytanie, jakie są czynniki różnicujące ich zachowania, inaczej — o dodatkowe zmienne warunkujące wpływ zmiennej niezależnej.

Każda z tych sytuacji jest dostatecznym powodem podjęcia nowych badań, teraz jednak ich celem nadrzędnym winno być sprawdzenie postawionych hipotez i zbadanie odkrytych zjawisk. Ale ponieważ okoliczności, w jakich je zaobserwowano były kontrolowane, to hipotezy te mają już za sobą godne uwagi — a niekiedy nawet wystarczające do ich przyjęcia — świadectwa empiryczne.

Eksperyment bywa też narzędziem diagnozy. Diagnoza to poznawanie ukrytych własności przedmiotu z ich objawów na podstawie uogólnień wiążących własności z objawami.

Własności ukryte, takie jak dyspozycje, potrzeby czy cechy osobowości, nie są z definicji dostępne bezpośredniej obserwacji badacza — może on tylko inferować je z ich obserwowalnych objawów lub — niekiedy — oprzeć się na introspekcyjnych zeznaniach badanych. Jeśli jednak rodzaj badanych przedmiotów i rodzaj cech ukrytych wykluczają tę drugą możliwość a objawów naturalnych brak lub też nie dają się one jednoznacznie interpretować, to badacz może sprowokować badane przedmioty do zachowań diagnostycznych i z tych zachowań wnioskować o cechach ukrytych. Podstawą projektowania eksperymentu i diagnozy jest twierdzenie ogólne mówiące, że jeśli badane przedmioty posiadają pewną własność, to w określonych warunkach zachowują się w pewien określony sposób. Rzeczą eksperymentu jest stworzyć owe warunki.

Tą drogą można stwierdzić np., czy badani mają postawy przesądne wobec grup etnicznych i czy są zatem etnocentrykami. Wystarczy w tym celu wypowiadać w ich obecności pozytywne opinie o nisko cenionych grupach etnicznych i mierzyć siłę ich odruchu skórno-galwanicznego Wiadomo bowiem, że etnocentrycy na opisaną sytuację bodźcową reagują wzmożoną reakcją emocjonalną i — dalej — wzmocnionym odruchem skórno-galwanicznym. W eksperymencie tym nie ustala się owej zależności, lecz się ją zakłada.

Można powiedzieć, że eksperymenty diagnostyczne zmuszają ukrytą rzeczywistość do odsłonięcia się.

Eksperymentuje się też po to, by zbadać i usprawnić istniejące oraz skonstruować' nowe techniki badawcze. Takie eksperymenty nazywa się metodologicznymi. Zmienia się w nich pewne elementy postępowania badawczego z zamiarem określenia ich wpływu na uzyskane wyniki a innowacje metodologiczne sprawdza się przez ich zastosowanie. Eksperyment — metoda badania staje się sam jego przedmiotem.

Im bardziej nauka staje się świadoma niedoskonałości swych metod, tym częściej badania tego typu się pojawiają — nauki społeczne nie są tu wyjątkiem. Szeroki nurt badań eksperymentalnych nad eksperymentami psychospołecznymi, zapoczątkowany w latach 60-tych w Stanach Zjednoczonych jest dobrym tego przykładem. W badaniach tych chodzi o ocenę wpływu rozmaitych aspektów społecznej sytuacji badania eksperymentalnego na jego przebieg i wyniki oraz o obmyślenie i sprawdzenie takich schematów i procedur eksperymentalnych, które zapewniałyby wynikom maksymalną trafność. Aczkolwiek bezpośrednim celem tych eksperymentów jest — tak jak i we wszystkich eksperymentach — obserwacja reakcji badanych na różnicowane warunki eksperymentalne, to celem nadrzędnym nie jest rozwój teorii, lecz metodologii; rzecz jasna metodologię rozwija się po to, by rozwijać teorię, ale to już w następnych badaniach.

Siła eksperymentu polega jednak, przynajmniej w socjologii; na sprawdzaniu hipotez a nie na heurystycznych i diagnostycznych użytkach czynionych z tej metody. Poznawaniu stanów rzeczy i stawianiu hipotez lepiej służą inne, mniej rygorystyczne a bardziej swobodne metody. Eksperyment nie jest więc wzorem badań empirycznych *tout court,* lecz tylko wzorem badań przyczynowych.

2.3. SZCZEGÓLNE ZALETY EKSPERYMENTU

Bez wątpienia najpoważniejszą zaletą eksperymentu jest łatwość, z jaką pozwala on sprawdzać hipotezy przyczynowe; łatwość ta nie jest dana metodom nieeksperymentalnym. Na eksperyment warto jednak spojrzeć i z innych punktów widzenia. Obraz jego stanie się wówczas pełniejszy a osobliwości — wyraźniejsze.

Eksperyment jest wzbogaceniem doświadczenia. Uwalnia badacza od skrępowania go faktami zastanymi. gdyż tworzy sytuacje nowe i zwielokrotnia rzadkie. Co więcej, zastany repertuar zmiennych, ich wartości i ich konfiguracji powiększa się w kierunkach nieprzypadkowych, lecz zgodnych z logiką badania.

Rzeczywistość naturalna jest skomplikowana, ale jej teorie są zwykle proste, sformułowane dla niewielkiej liczby zmiennych przy świadomym wyłączeniu innych. Eksperyment przybliża rzeczywistość do modelu teorii, dziedziny, w której jest ona zinterpretowana. Związki istniejące w rzeczywistości ukazane są przez to w stanie czystym, wolne od zakłóceń, łatwe do obserwacji.

W innych wypadkach logika badania wymaga większego zróżnicowania wartości zmiennych niż zróżnicowanie obserwowane w warunkach naturalnych. Prawa ilościowego nie potrafimy sprawdzić, jeśli zmienna zależna przybiera wartości nie reprezentujące całej jej skali a tylko np. środek *continuum.* Eksperyment pozwoli wówczas uzupełnić zmienność badanej zmiennej o wartości skrajne.

Jeszcze innym razem logika badania wymaga dostępu do wszelkich możliwych kombinacji kilku zmiennych. Jest tak wtedy, gdy bada się efekty interakcji zmiennych. Jeśli takich kombinacji nie ma w warunkach naturalnych, a zwykle nie ma, to można je sobie stworzyć. Eksperyment czynnikowy na tym właśnie polega.

Eksperyment to pytanie skierowane do natury — mówi popularna metafora. Można ją rozwijać: samo tylko podglądanie rzeczywistości dostarczyłoby danych tak ubogich, jak zastąpienie badań ankietowych podsłuchiwaniem.

Eksperyment podwyższa jakość obserwacji i opisu zjawisk. Ponieważ jest doświadczeniem uprzednio zaplanowanym, świadomym i celowym wywoływaniem zjawisk, to badacz nie jest ich pojawieniem się zaskoczony — on ich oczekuje. Może więc przygotować się do obserwacji i rejestracji i prowadzić je skrupulatniej. Psychologia postrzegania poucza, jak znaczne są różnice w ilości a szczególnie w rzetelności informacji uzyskanych w trakcie obserwacji zjawisk z jednej strony oczekiwanych, a z drugiej — niespodziewanych. Poucza również, że w obserwacji tych drugich różnice wyników powodowani indywidualnymi cechami obserwatorów są szczególnie duże. Eksperyment zwiększa więc ilość i jakość uzyskanych danych, a zmniejsza różnice wywołane odmienną percepcją różnych badaczy. Eksperymentator ponadto posługuje się zazwyczaj specjalnymi instrumentami pomiarowymi, które pozwalają mu dotrzeć do ukrytych aspektów zjawisk, zmierzyć je z Większą dokładnością i pewnością. Nie musi on polegać wyłącznie na swych zmysłach. W eksperymencie kontroluje się więc nie tylko działanie bodźców, lecz także obserwację ich następstw.

Eksperyment w szczególnie przemyślany sposób wiąże teorię z danymi empirycznymi, rozum z doświadczeniem.

Eksperyment wykonuje się po to, by odpowiedzieć na pewne pytanie — inaczej nie byłoby wiadomo, dlaczego w ogóle został on przeprowadzony. Problemy rozwiązywane za pomocą eksperymentu wyrastają z pewnego stanu teorii naukowej a ich rozwiązania dają ’początek nowym problemom. Ale pytania, na które eksperyment odpowiada, to nie są byle jakie pytania. Są to pytania rozstrzygnięcia a formułowanie takich pytań wymaga dużej wiedzy przedmiotowej. Gdyby eksperymentator nie kierował się hipotezami — próbnymi odpowiedziami na owe pytania, to eksperyment stałby się pracochłonną i nieefektywną metodą prób i błędów, której stosowanie przypadkowo tylko mogłoby doprowadzić do wartościowych wyników.

Przekład figurujących w sprawdzanych hipotezach zmiennych niezależnych na manipulacje eksperymentalne również wymaga wiedzy teoretycznej. Bodźce eksperymentalne a często i cały kontekst badania zmierzają do zrealizowania warunków określonych w poprzedniku hipotezy. Gdyby nie ona, nie byłoby wiadomo, dlaczego wprowadza się bodźce takie a nie inne. Przyjmuje się także założenie o skuteczności manipulacji. Założenie to pochodzi spoza eksperymentu i nie zawsze jest prawdziwe. Eksperymenty dają niekiedy wyniki sprzeczne z hipotezą nie dlatego, że jest ona fałszywa, lecz dlatego, że. fałszywe było założenie o skuteczności manipulacji.

Również plan kontroli i schemat eksperymentu określone są przez pochodzącą skądinąd wiedzę teoretyczną. Tak kontrola przez izolację bodźca od jego zakłóceń, jak i przez grupy kontrolne wymaga uprzedniego odróżnienia zmiennych istotnych, które trzeba izolować lub wyrównać od zmiennych nieistotnych, które można pominąć. Podział taki wymaga odwołania się do twierdzeń przyczynowych o istnieniu bądź nieistnieniu związków przyczynowych między zmiennymi rozważanymi a zmienną zależną. Reszta jest sprawą logiki i wnioskowań statystycznych. Ich wartość zależy jednak od zasadności przyjętych założeń o istotności a zwłaszcza o nieistotności zmiennych towarzyszących bodźcom. Istnieje możliwość krytyki eksperymentu przez kwestionowanie tych założeń i wykazywanie, że pominięto pewne czynniki istotne, bowiem błędnie uznano je za nieistotne. Taka krytyka równa się wysunięciu konkurencyjnego względem sprawdzanej hipotezy wyjaśnienia różnic między grupą eksperymentalną i kontrolną. W gruncie rzeczy kwestionuje ona wartość całego eksperymentu.

Wreszcie, pomiar zmiennej zależnej opiera się na pewnych założeniach teoretycznych. Odnoszenie zmian zaobserwowanych do zmian w wartości zmiennej teoretycznej wymaga uprzedniego przyjęcia twierdzeń o zależnościach między jednymi a drugimi. Jest to przypadek znanego zagadnienia operacjonalizacji zmiennych teoretycznych. Jeśli pomiar tych zmiennych- odbywa się przy pomocy specjalnych instrumentów, to dyskutowane założenia przybierają formę teorii, na której oparto konstrukcję narzędzia oraz założeń o jego rzetelności. Tak mierzenie temperatury ciała termometrem rtęciowym zakłada stwierdzenie o rozszerzaniu się rtęci pod wpływem ciepła a pomiar nieuświadomionych lub ukrywanych dyspozycji przy pomocy testu projekcyjnego zakłada psychoanalityczną teorię projekcji. O obu narzędziach zakłada się ponadto, że ich wrażliwość na rozmaite okoliczności zewnętrzne nie przekracza granic tolerancji.

Z drugiej strony, eksperyment pomaga rozwijać teorię w sposób zdyscyplinowany.

Eksperyment umożliwia empiryczną interpretację takich zmiennych teoretycznych, które w warunkach naturalnych byłoby trudno lub nie sposób zoperacjonalizować; gdyby nie eksperyment, to teorie, w których takie zmienne figurują, pozostałyby co najwyżej w zasadzie sprawdzalnymi. Dalej, eksperyment umożliwia sprawdzanie teorii idealizacyjnych, sformułowanych dla pewnej niewielkiej liczby zmiennych a pomijających inne; gdy takich możliwości nie ma, to teorie posądzane są o spekulatywność. Wreszcie, operacjonalizacja terminów teoretycznych na potrzeby eksperymentu jest pod pewnym względem szczególnie wartościowa; trzeba praktycznie wytworzyć ich obserwacyjne odpowiedniki; nad trudnościami pojęciowymi możemy się prześlizgiwać dopóty, dopóki manipulujemy tylko słowami — zauważa Kelman[12].

Z uwagi na duży i wyraźny udział myślenia teoretycznego w projektowaniu i prowadzeniu eksperymentów, metodologiczna analiza eksperymentu jest obfitym źródłem argumentów w dyskusjach nad stosunkiem faktów i teorii w procesie rozwoju nauki. Z tego, co powiedziano wyżej wynika, że takie zalecenia, jak „Obserwuj i uogólniaj wyniki obserwacji! Najpierw obserwacja, potem teoria! Obserwuj nie sugerując się żadną teorią!" nie dadzą się pogodzić z ideą i praktyką eksperymentowania. Nie można więc skorzystać z czyjejś rady, by wchodząc do laboratorium, zostawiać za progiem własne teorie, tak jak zostawia się płaszcz w szatni. Żeby coś znaleźć, trzeba najpierw wiedzieć, czego się szuka. Inna rzecz, że posiadane przez eksperymentatora hipotezy mogą go uczynić ślepym na fakty świadczące, przeciw nim. Zapobieganiu temu służy powtarzanie eksperymentu.

Eksperyment w szczególnie efektywny sposób czyni badania naukowe intersubiektywnymi. Eksperymentator manipulujący warunkami zachodzenia i przebiegu zjawisk może warunki te opisać' w sposób przyjęty w języku społeczności naukowej. Dzięki temu umożliwia wielokrotne powtórzenie eksperymentu i sprawdzenie jego wyników sobie i — co ważniejsze — innym badaczom dysponującym

[12] H. C. Kelman, Time to speak: on human values and social research, San Francisco 1968.

stosowną kompetencją i środkami. Każdy eksperyment może być powtórzony a jego wynik może okazać się niezgodny z wynikiem pierwotnym. Niebezpieczeństwo to zmusza badaczy do samokontroli i do przeprowadzania eksperymentów w tak poprawny sposób, by nie budził on zastrzeżeń i chęci powtórzenia całego badania. Eksperymentatorzy stają się sami swymi krytykami. Powtarzalność eksperymentu jest zatem sprężyną w mechanizmie samokontroli działającym tak na poziomie poszczególnych badaczy, jak też na poziomie nauki jako całości. Przyczynia się ona do tego, że nauka jest zbiorowym przedsięwzięciem i zbiorowym dobrem.

3. KRÓTKA HISTORIA EKSPERYMENTU

Omówione wyżej zalety, właściwe tylko eksperymentowi lub też właściwe mu w szczególnym stopniu, uczyniły metodę eksperymentalną najpotężniejszą, obok abstrakcji teoretycznej, dźwignią rozwoju nowożytnych nauk przyrodniczych.

Pierwszą nauką, która skorzystała z dobrodziejstw eksperymentu był fizyka. W XVII wieku nauka ta zerwała z naiwną indukcją, opartą na doświadczeniu potocznym i zdrowym rozsądku. Abstrakcja pozwoliła jej budować hipotetyczne modele zjawisk, uwzględniające tylko wybrane czynniki i ustalające proste zależności między nimi, ujęte w formuły matematyczne. Eksperyment umożliwił sprawdzenie w uproszczonych i dobrze kontrolowanych warunkach tych konsekwencji owych modeli, które się do tego nadawały. Zarówno abstrakcję teoretyczną jak i eksperyment stosowano sporadycznie już w starożytności i wiekach średnich, jednakże wówczas ani abstrakcja nie była kontrolowana, ani eksperyment nie służył sprawdzaniu teorii, więc i efekty ich nie były wielkie. Natomiast metoda hipotetyczno-eksperymentalna stanowiła mariaż teorii i eksperymentu. Skojarzył ten związek Galileusz a jego owocem było gwałtowne przyśpieszenie postępu nauki w XVII wieku[13]. Uczeni ówcześni doskonale zdawali sobie sprawę z wartości nowego podejścia: „Sądzę, że nic nie przyczynia się tak bardzo do postępu filozofii [naturalnej], jak sprawdzanie hipotez przez eksperymenty i dochodzenie do eksperymentów przez hipotezy” — napisał w 1672 r. Robert Hooke[14].

Metoda eksperymentalna, zastosowana z takim powodzeniem w fizyce, odegrała również znakomitą rolę w rozwoju innych nauk przyrodniczych.

W chemii Lavoisier w drugiej połowie XVIII wieku pierwszy zastosował na szeroką skalę eksperyment pomyślany jako czynność sprawdzająca uprzednio postawione hipotezy, w dodatku oparta na pomiarach i kontrolowana przez wielokrotne powtarzanie. Dawni chemicy zdawali się raczej na przypadek. Choć również eksperymentowali, to zwykle po to, by zobaczyć, jak zachowa się dany układ w warunkach dobranych przypadkowo; uważali, że przyjęcie , z góry hipotez jest szkodliwe. Lavoisier odkrył dla chemii eksperyment w sensie nowoczesnym, sterowany przez hipotezy. Odkryciu temu zawdzięczamy w dużej mierze rewolucję chemiczną z teorią tlenową jako jej głównym elementem i chemię w jej nowoczesnej postaci[15].

Claude Bernard dowiódł teoretycznie, że eksperyment stosowany dotąd tylko w fizyce i chemii można wykorzystać także w naukach biologicznych. Uzasadnił on tezę o jedności przyrody żywej i martwej, pokazał, że złożona organizacja i spontaniczność zachowania organizmów żywych nie uniemożliwia eksperymentowania z nimi a tylko je utrudnia, gdyż środowisko wewnętrzne, dla nich specyficzne, podporządkowane jest zasadzie determinizmu, tak samo jak właściwe rzeczom

[13] Zob. A. C. Crombie, Nauka średniowieczna i początki nauki nowożytnej, Warszawa 1960, t. 1, cz. II, rozdz. 1; R. Hooykaas, Religia i powstanie nowożytnej nauki, Warszawa 1975

[14] R. Hooke, List do Lorda Brouncker (?), w: H. W. Turnbull (ed.), The correspondence of Isaac Newton, Cambridge 1959, vol. 1, s. 202.

[15] Zob. E. Pietruska-Madej, Metodologiczne problemy rewolucji chemicznej, Warszawa 197b, s. 115 - 127.

martwym środowisko zewnętrzne. Słuszności tych idei dowiódł praktycznie znakomitymi wynikami „medycyny eksperymentalnej". Nauka ta oparta była głównie na eksperymentach ze zwierzętami, lecz jej wyniki odnosiły się także do ludzi[16].

Zdumiewające postępy nauk eksperymentalnych — fizyki, chemii, biologii stały się glebą, na której wyrósł mit eksperymentu — jedno z wcieleń dobrze znanej historykom nauki skłonności do utożsamiania jakiegoś jednego owocnego w pewnych dziedzinach podejścia z metodą naukową *tout court.* Tendencja do stawiania znaku równoważności między eksperymentalną a metodą naukową przyjęła dwa oblicza.

Z jednej strony wyrażała się ona w niezwykle pozytywnym wartościowaniu metody eksperymentalnej. Uznawano ją za „prawdziwą" metodę naukową, najlepiej wyrażającą empiryzm nauki i realizującą jej zalania. Wyniki badań eksperymentalnych obdarzono zaufaniem *a priori.* Utrzymywano, że eksperymentowanie jest jedyną drogą unaukowienia niedostateczne rozwiniętych dyscyplin. Dzielono historię nauk na okresy przedeksperymentalne i eksperymentalne. Czyniono nauki eksperymentalne wzorem służących do oceny innych nauk i punktem, ku któremu winien zmierzać ich rozwój.

Z drugiej strony, mit eksperymentu znajdował wyraz w skłonności do lekceważenia badań prowadzonych innymi metodami oraz do odrzucania teorii opartych na innych badaniach niż eksperymentalne, przy czym postawa ta nie była dyktowana przekonaniem o zawodności owych metod i fałszywości teorii, lecz przekonaniem o ich nienaukowości. Wiadomo np., że opór przeciw biologicznej teorii ewolucji był uzasadniany również brakiem jej potwierdzenia eksperymentalnego[17]. Istotnie, Darwin nie odtwarzał i nie obserwował procesów doboru naturalnego w laboratorium, lecz tylko rekonstruował je w sposób pośredni z danych historycznych i istniejącego zróżnicowana organizmów. Okoliczność ta sama w sobie jest wszakże bez znaczenia dla oceny wartości jego teorii.

Bezzasadność dyskutowanej równoważności jest zbyt oczywista, by się nad nią dłużej zatrzymywać. Wielu filozofów nauki przekonująco uzasadniło pogląd, że eksperyment nie jest warunkiem koniecznym uzyskania wiedzy godnej zaufania — warunkiem tym jest natomiast kontrolowany charakter badań, który można zapewnić w wielu dziedzinach bez posługiwania się manipulacją eksperymentalną. Badania metodologiczne wykazały też, że eksperyment sam w sobie nie zapewnia wiedzy rzetelnej — wymaga on rozwiniętej teorii, pomysłowej manipulacji, dobrych narzędzi pomiarowych. Przekonania te należą obecnie do sądów obiegowych, żeby nie powiedzieć — banalnych.

Niemniej jednak eksperyment pozostaje narzędziem w wielu dziedzinach badań podstawowym lub wręcz jedynym, a w wielu innych nad wyraz cenionym. Nawet przedstawiciele dyscyplin zasadniczo nieeksperymentalnych, takich jak historia czy ekonomia polityczna, lubią czasem ponarzekać na to, że nauki ich są takimi właśnie. Przy całym szacunku dla innych rodzajów świadectw uważa się, że jeśli

[16] Zob. C. Bernard, Wstęp do medycyny eksperymentalnej, w: B. Skarga, Claude Bernard, Warszawa 1970; J. Schiller, The genesis and structure of Claude Bernard's experimental method, w: R.N. Giere, R. S. Westfall (eds.), Foundations of scientific method: the nineteenth century, Bloomington 1973.

Ciekawe, że sukcesy eksperymentalnego badania zjawisk życia stały się tłem koncepcji uprawiania literatury pięknej na podobieństwo badań i raportów eksperymentalnych. Zafascynowany pracami Bernarda teoretyk naturalizmu, Emil Zola, poszedł jego drogą krok dalej. Rozszerzywszy założenie determinizmu na życie psychiczne i społeczne człowieka, Zola proponuje eksperymentowanie na jego emocjach i intelekcie — literatura powinna przyjąć na siebie to zadanie: „powieść naturalistyczne […] jest prawdziwym eksperymentem, jaki pisarz wykonuje ze wsparciem obserwacji na ludziach", zastąpi ona „powieść czystej wyobraźni". Pisarz — naturalista bierze fakty z natury i ustala sytuację wyjściową, po czym wprawia je w ruch: modyfikuje, stawia w nowych warunkach, poddaje testom nowych kontrolowanych sytuacji, aby sprawdzić i pokazać, jak w tych sytuacjach działają emocje i intelekt; wszystko to w zgodzie z deterministycznymi prawami natury. Wynikiem jest „naukowa wiedza o człowieku". „Powieść eksperymentalna — pisze Zola — jest po prostu sprawozdaniem z eksperymentu, jaki pisarz powtarza na oczach publiczności". Jest to ważne świadectwo tryumfów nauk eksperymentalnych a i światopoglądu scjentystycznego również. Zob. E. Zola, The experimental novel (1880), w: R. Olson (ed.), Science as metaphor, Belmont 1971, cyt. ze s. 237.

[17] Podaję za: Th. Dobzhansky, Adaptive changes induced by natural selection in wild population of Drosophila, w: M. L. Gabriel, S. Fogel (eds.), Great experiments in biology, Englewood Cliffs 1955.

jakaś teoria, dotąd uzasadniana nieeksperymentalnie, zyskuje potwierdzenie eksperymentalne, to stopień jej potwierdzenia znacząco się podnosi i zbliża się ona do jednomyślnego uznania przez społeczność naukową. Prowadzi to do szukania możliwości eksperymentu w nieraz zgoła nieoczekiwanych dziedzinach badań.

Przykładem niech będą losy teorii Oparina, wedle której życie na Ziemi powstało samorzutnie ze związków nieorganicznych. Teorię tę uzasadniano, bez większego zresztą powodzenia, tylko za pomocą innych teorii aż do lat pięćdziesiątych, kiedy to Miller wymyślił dla niej test eksperymentalny. Zbudował on model protooceanu i protoatmosfery, który zawierał te związki organiczne, które się w nich prawdopodobnie znajdowały. W warunkach silnych wyładowań elektrycznych, imitujących światło słoneczne o wysokiej częstotliwości, ze związków nieorganicznych powstały substancje organiczne — proste składniki organizmów żywych. Główna trudność teorii tj. sama możliwość takiej reakcji, została więc usunięta eksperymentalnie[18]

Pierwszym eksperymentatorem w dziedzinie obejmowanej dziś przez nauki społeczne był może Psammetych, faraon panujący w VII wieku p.n.e. „Kiedy […] Psammetych wstąpił na tron i zapragnął poznać, którzy ludzie naprzód zostali stworzeni […] — pisze Herodot — wpadł na taki pomysł. Oddał dwoje nowo narodzonych dzieci z pierwszych lepszych rodziców pasterzowi do jego trzody z poleceniem, by je tak wychowywał, żeby nikt w ich obecności żadnego głosu z siebie nie wydał […]. To uczynił i polecił Psammetychus, ponieważ chciał wiedzieć, jaki dźwięk pierwszy po niewyraźnym kwileniu dzieci z siebie wydobędą. Tak się też stało”[19]. Pomysł faraona można uznać za pierwszą znaną próbę rozwiązania w drodze eksperymentalnej niezwykle żywego w dziejach nauk społecznych zagadnienia: jakie cechy wykształciłaby w sobie jednostka dojrzała poza społeczeństwem?

Przypadek Psammetycha ma charakter raczej tylko anegdotyczny, a eksperymentalne pasje faraona nie znalazły naśladowców w kolejnych pokoleniach badaczy spraw ludzkich[20]. Systematyczne eksperymentowanie na ludziach podjęto dopiero w drugiej połowie XIX wieku. Myśl zastosowania metody eksperymentalnej w „naukach politycznych i moralnych” rzucił jako pierwszy bodaj Laplace w r. 1814 — oczekiwał on, że eksperymentalne badanie zagadnień społecznych doprowadzi te nauki do wyników tak świetnych, jak osiągnięcia eksperymentalnych nauk przyrodniczych. Możliwości eksperymentu w naukach społecznych rozważał szeroko w opublikowanym w 1843 r. Systemie logiki John Stuart Mill. Doszedł on jednak do wniosku, że eksperyment w tych naukach nie jest możliwy, gdyż między zjawiskami społecznymi istnieje zbyt wiele różnic i zbyt wiele podobieństw, by móc zastosować do ich analizy jego kanony eksperymentalne[21]

[18] Zob. M.G. Rutten, Powstanie życia na Ziemi a świadectwa geologiczne, Warszawa 1966, s. 65-71.

[19] Herodot, Dzieje, Warszawa 1954, s. 119-20.

Inny efektowny przypadek opowiada Ateneusz, pisarz grecki z II w. n.e. Opisuje on, jak odkryto eksperymentalnie antytoksyczne właściwości cytronu: Rzecz dziać się miała w Egipcie. Grupę przestępców skazano na śmierć przez ukąszenie żmii. Mimo ukąszeń, pozostali oni przy życiu. Zarządzone w tej sprawie śledztwo wykazało, że prowadzonym na śmierć litościwa straganiarka podała cytrony. Aby sprawdzić, czy to cytrony uratowały im życie „nakazano dać kawałek cytronu, dokładnie jak przedtem, jednemu skazańcowi, lecz nie drugiemu i ten, który go zjadł, nie doznał żadnej krzywdy po ukąszeniu przez gady, drugi umarł w chwilę po ugryzieniu go. Tak więc, ponieważ ten sam wynik uzyskano w wielu przypadkach, dowiedzione zostało, że cytron jest antidotum na każdą (sic!) truciznę”. (Sofiści przy uczcie, ks. III, 84-5; korzystaliśmy z przekładu ang.: The deipnosophists, Cambridge, Mass. 1951).

[20] Z jednym wyjątkiem. Fryderyk II, cesarz niemiecki i król jerozolimski (1194- 1240), znany historykom nauki ze swych pasji ornitologicznych, wykonał był także parę, mniej znanych, eksperymentów z zakresu nauk o człowieku. W jednym z nich, chcąc sprawdzić czy człowiek ma duszę, kazał zamknąć pewnego nieszczęśnika w szczelnie zapieczętowanej beczce, a gdy po jej otworzeniu znalazł w niej tylko zwłoki, doszedł do wniosku, że człowiek duszy nie ma. W innym doświadczeniu poszedł on dokładnie — zapewne bezwiednie — w ślady Psammetycha. Zabronił matkom nowo narodzonych dzieci mówić do nich, gdyż chciał zobaczyć, czy dzieci zaczną z czasem mówić językiem swych rodziców czy też może jakimś starszym językiem, np. hebrajskim czy greckim. Nie wiemy, do jakich wniosków doszedł cesarz tym razem, zauważył on jednakże, iż niemowlęta źle się chowały bez wyrazów czułości. Tak więc, odkrył on przypadkowo rolę matczynej miłości w rozwoju - dzieci, i to na 7 wieków przed Harlowem. Wiadomości te podajemy za: R.E. Juday, Versatile genius: Frederick II, „Science’’ 1968, vol. 162, nr 3857, s. 850 - 1.

[21] Zob. J.S. Mili, System logiki, Warszawa 1962, t. 1, ks. III, rozdz. 7.

Niemniej jednak, niezależnie od teoretycznych dyskusji nad możliwością stosowania eksperymentu w naukach społecznych, metoda ta zaczęła powoli zdobywać miejsce w instrumentarium nauk społecznych, przesądzając tym samym wynik owych debat.

Zaczęło się od psychologii[22], gdzie Weber i Fechner podjęli w drugiej połowie XIX wieku rozległe badania z zakresu psychofizyki i psychofizjologii. Skupiały się one głównie na zależnościach wrażeń od pewnych cech bodźców fizycznych działających na różne zmysły; próg wrażeń czy czas reakcji były typowymi problemami tego okresu. To, że eksperymentować zaczęto najpierw na pograniczu psychologii i eksperymentalnych nauk przyrodniczych dziwić nie może. Stopniowo jednak dziedzina eksperymentowania rozszerza się i obejmuje zjawiska bardziej odległe od psychofizyki i psychofizjologii, takie jak pamięć czy emocje. W r. 1879 Wundt zakłada w Lipsku pierwsze badawcze laboratorium psychologiczne, a podobne pracownie powstają niebawem w innych miejscach.

W laboratoriach tych zaczyna się powoli uwzględniać zmienne społeczne, zrazu — rzecz zrozumiała — mało skomplikowane. Allport[23] zbadał, że przez trzy dziesięciolecia aż do roku mniej więcej 1920 psychologowie zajmowali się właściwie tylko jedną zmienną społeczną a mianowicie obecnością innych osób przy wykonywaniu przez pojedynczych ludzi pewnych czynności i starali się ustalić, jaki wpływ ma ta zmienna na wykonywanie czynności. W typowym dla tego okresu eksperymencie, Triplett ustalił w 1898 r., że badani wykonywali pewną prostą czynność szybciej, gdy pracowali w towarzystwie innych, robiących to samo osób; zjawisko to znane jest obecnie jako efekt facylitacji społecznej. Grupy społeczne spotykane w ówczesnych eksperymentach były tylko zbiorowościami osób działającymi razem *(coacting groups).*

W latach dwudziestych wyróżniono obok nich grupy rozumiane jako zespoły osób pozostających we względnie trwałych bezpośrednich interakcjach *(face-to--face groups).* Otwarło to przed badaczami nowe problemy teoretyczne i empiryczne, związane ze wzorami interakcji, strukturą, celami, normami grup. Grupy tego rodzaju stały się właściwym przedmiotem socjologii małych grup czyli mikrosocjologii. Podstawową metodą tej dyscypliny stał się eksperyment. W okresie międzywojennym wykonano wiele eksperymentów **z** zakresu socjologii małych grup. Z najwcześniejszych można wymienić przeprowadzone w latach dwudziestych w ZSRR — nie bez wpływu nowej ideologii — eksperymenty Bechterewa i de Lange, w których porównywano zachowania indywidualne z kolektywnymi np. efektywność myślenia indywidualnego i grupowego, współpracy i konkurencji. Innym, bardziej znanym przykładem są wykonane w latach trzydziestych eksperymenty Sherifa nad rozwiązywaniem konfliktów międzygrupowych[24]. Z tych lat pochodzą też inne eksperymenty terenowe. W latach 1930—1933 Dodd wykonał eksperyment terenowy w Syrii, przy pomocy którego badał wpływ szkolenia w zakresie higieny na poziom higieny na wsi[25]. Jeszcze wcześniej, bo w r. 1924 rozpoczęły się w zakładach Hawthorne słynne badania, których efektem teoretycznym było odrzucenie taylorowskiej teorii organizacji na rzecz teorii stosunków międzyludzkich, a efektem metodologicznym odkrycie wpływu samego eksperymentowania na zachowanie, nazwanego potem efektem Hawthorne[26].

Największe znaczenie dla rozwoju eksperymentowania w socjologii miała jednak działalność Kurta Lewina[27]. Lewin pierwszy przedstawił pod koniec lat trzydziestych szeroki program eksperymentowania w „przestrzeni społecznej" i gło-

[22] Zob. E. G. Boring, A history of experimental psychology, New York 1950.

[23] Zob. G.W. Allport, Historical background of modern social psychology, w: G. Lindzey, E. Aronson (eds.), The handbook of social psychology, Reading, Mass. 1968, vol. 1.

[24] Zob. Th. M. Newcomb i in., Psychologia społeczna, Warszawa 1970, s. 481 - 3.

[25] Zob. F.S. Chapin, Experimental designs in sociological research, New York 1955, s. 55 - 7.

[26] E. L. Cass, F. G. Zimmer (eds.), Man and work in society. A report on the symposium held on the occasion of the 50 th anniversary of the original Hawthorne studies…, New York 1975, zob. też przyp. 8 na s. 36.

[27] Zob. D. Martindale, The nature and types of sociological theory, Boston, 1960, s. 50.3 - 12.

sił, że eksperyment w socjologii jest możliwy tak samo jak w fizyce czy chemii. Ośrodkiem eksperymentowania miała stać się w jego koncepcji grupa, a zadanie eksperymentatora miało polegać na tworzeniu grup, ich klimatów i stylów życia.

Ważniejsze niż samo sformułowanie programu było jednak poparcie go wieloma eksperymentami nad życiem grupy, a zwłaszcza nad wpływem jej klimatu na zachowanie członków. Najbardziej znanymi z tych badań są eksperymenty nad wpływem demokratycznej, autorytarnej i liberalnej władzy w grupie na jej życie. Badania grupy Lewina otwarły szeroko wrota dla eksperymentalnego badania różnych aspektów życia grupy: formowania się grup i ich spójności, nacisków i norm grupowych, celów grupowych i ich osiągania, struktury i przywództwa, konfliktów wewnątrz- i międzygrupowych itd. Skutki tych badań dla zrozumienia podstawowych mechanizmów życia społecznego, stosunkowo łatwo dających się obserwować w małych grupach, okazały się zachwycające. Socjologia małych grup- jest dziś może najbardziej rozwiniętym działem socjologii, jeśli za miarę rozwoju nauki wziąć zasób systematycznie powiązanych, porządnie skontrolowanych twierdzeń, zwłaszcza zaś twierdzeń ogólnych.

Śmiałe stosowanie eksperymentu walnie przyczyniło się do szybkiego rozwoju najpierw fizyki a potem kolejno chemii, biologii, psychologii a wreszcie mikrosocjologii, nauki wyrosłej na pograniczu psychologii z socjologią. Czy metoda eksperymentalna może odegrać tę samą rolę także w odniesieniu do tych dziedzin socjologii, które nie ograniczają swych zainteresowań do małych grup, lecz uwzględniają wyższe niż mała grupa poziomy analizy, bardziej skomplikowane interakcje i bardziej zróżnicowane warunki. Czy eksperyment stać się może narzędziem makrosocjologii? Bez wątpienia, nie. Nawet sama nazwa „socjologia eksperymentalna" pojawia się bardzo rzadko i brzmi dziwnie. Jednakże w ostatnich latach, w wielu krajach, wraz z nasileniem się tendencji do naukowego planowania zmiany społecznej odżyła ciekawa idea „społeczeństwa eksperymentującego"[28], rzucona przed trzydziestu laty przez Skinnera w powieści *Walden Two* (1948)[29]. Idea ta postuluje wypróbowanie w drodze eksperymentów *in vivo* nasuwających się rozwiązań problemów społecznych, ocenianie ich w sposób nieintuicyjny i wielopłaszczyznowy, by poprzez eliminację propozycji wadliwych dochodzić do przyjęcia rozwiązań najlepszych. Co ważniejsze, eksperymentów takich prowadzi się coraz więcej, a metodologia eksperymentalnej ewaluacji programów społecznych rozwija się szybko.

Tak oto metoda eksperymentalna przebyła drogę po szczeblach drabiny nauk od fizyki do socjologii.

Ciekawe, że wraz z inwazją eksperymentu na teren nauk społecznych przeniosła się również omawiana wyżej skłonność do fetyszyzacji tej metody. Tym razem miała ona za podstawę już nie tylko rolę eksperymentu w rozwoju przyrodoznawstwa, lecz także sukcesy eksperymentalnych badań społecznych. Błyskotliwy krytyk stanu nauk o zachowaniu sprzed lat kilkunastu pisał: „We współczesnej nauce o zachowaniu postawie wobec eksperymentu grozi to, że stanie się ona rodzajem rytuału: jak gdyby samo przyłożenie rąk mogło spowodować ozdrowienie chorych idei. Jak w każdym rytuale, nacisk przenosi się z treści na formę, z kwestii merytorycznych na proceduralne, a cnotą staje się wykonanie skończonego ciągu czynności. Poszczególne techniki utożsamia się z «metodą naukową», a badania, w których nie używa się ich. lekceważy się wówczas jako nie mające znaczenia naukowego"[30].

Mniej więcej w tym samym czasie inni badacze amerykańscy Hartley'owie zauważyli, że „metoda eksperymentalna ma tak silny wpływ na wyobraźnię publicz-

[28] Zob. D. T. Campbell, Reforms as Experiments, w: E. L. Struening, M. Guttentag (eds.), Handbook of evaluation re- search, Beverly Hills 1975, vol. 1; W. Jungk, Der Jahrtausendmensch. Bericht aus den Werkstätten der neuem Gesellschaft, München 1973, przekł. polski Człowiek tysiąclecia, w druku.

[29] O tym, że Skinner nie uważał swych pomysłów za literackie fantazje, lecz za program społeczny, świadczy jego wstęp do nowego wydania tej powieści (1976) oraz inna jego książka — Poza wolnością i godnością, Warszawa 1978.

[30] A. Kaplan, The conduct of inquiry, San Francisco 1964, s. 146.

ności, że istnieje skłonność do przyjmowania wyników za trafne właśnie dlatego, że zastosowano tę metodologię" a Stanisław Ossowski opatrzył ich uwagę komentarzem następującym: „Wydaje mi się, że uwaga Hartleyów [...] ma w obecnej chwili nie mniejszą aktualność w Polsce niż w Ameryce"[31].

W pracy tej będziemy mieli wiele okazji do spojrzenia chłodnym okiem na stosowalność i wartość eksperymentu w badaniach zjawisk społecznych — w istocie nie czemu innemu jest ona poświęcona.

[31] S. Ossowski, O osobliwościach nauk społecznych, Dzieła t. IV, Warszawa 1967, s. 278 - 279.

CZĘŚĆ II SCHEMATY I PROCEDURY EKSPERYMENTALNE W BADANIACH SPOŁECZNYCH

Eksperyment może służyć różnym celom, lecz zwykle służy sprawdzaniu hipotez przyczynowych. Aby cel ten osiągnąć, trzeba najpierw określić rodzaj manipulacji, po wykonaniu której wolno będzie powiedzieć, że został zrealizowany poprzednik sprawdzanej hipotezy oraz rodzaj zachowania, po zaobserwowaniu którego wolno będzie powiedzieć, że został zrealizowany również jej następnik. Jest to wymóg trafności teoretycznej. Dalej, trzeba tak zaplanować schemat eksperymentu i tak wykonać procedury eksperymentalne, by można było przypisać manipulacji wszystkie te i tylko te zachowania, które ona faktycznie spowodowała. Jest to wymóg trafności wewnętrznej. Są to podstawowe warunki dobrego eksperymentu — jeśli eksperyment ich nie spełni, to nie spełni i swego zadania.

W niniejszej części pracy rozważymy zagadnienie przekładu hipotez teoretycznych na hipotezy operacyjne, sformułowane w języku bodźców i reakcji eksperymentalnych. Potem omówimy czynniki uboczne, które mogą nakładać się na działanie bodźców i utrudniać ocenę ich rzeczywistego wpływu; czynniki te muszą więc być kontrolowane. To uczyniwszy, przystąpimy do przeglądu schematów eksperymentalnych pokazując przy tym, w jaki sposób pozwalają one badaczowi realizować postulat kontroli i przez to osiągać cele eksperymentowania. Oddzielny rozdział poświęcimy eksperymentowi społecznemu, ’mało znanej a — jak sądzimy — wielce obiecującej odmianie eksperymentu.

4. MANIPULACJA EKSPERYMENTALNA

Eksperymentalne sprawdzenie hipotezy przyczynowej wymaga najpierw trafnego przełożenia figurujących w niej zmiennych teoretycznych na język operacji badawczych i skutecznego wykonania tych operacji. Jeśli ten warunek nie zostanie spełniony, to jeśli nawet ustalimy ponad wszelką wątpliwość, że pewna manipulacja powoduje pewne zarejestrowane zmiany zachowania, nie będziemy mogli utrzymywać z równą pewnością, że potwierdziliśmy określoną hipotezę. Choć wynik eksperymentu będzie jednoznaczny na poziomie operacyjnym, to będzie on wieloznacznym na poziomie teoretycznym. Możliwości wielorakiej teoretycznej interpretacji zależności eksperymentalnej muszą więc być ograniczone tak bardzo, jak tylko jest to możliwe. W związku z tym przed eksperymentatorem stają trzy problemy:

1. Jakie zdarzenie uznać za operacyjny odpowiednik pewnego poziomu teoretycznej zmiennej niezależnej?
2. Jak zdarzenie to zrealizować, by wpłynąć na poziom zmiennej niezależnej?
3. Jak zmierzyć teoretyczną zmienną zależną?

Pytanie trzecie jest w tym miejscu najmniej ciekawe. W takim stopniu, w jakim dotyczy ono wszelkich pomiarów zmiennych teoretycznych, nie- interesuje nas ono. Natomiast o wpływie społecznych wymiarów sytuacji badawczej na wyniki po-

miaru będziemy mówić w osobnym rozdziale. Pozostaje się skupić na manipulacji — konstruowaniu i wprowadzaniu bodźca[32].

Zauważmy na początek, że w socjologii małych grup i psychologii społecznej niewiele jest standardowych sposobów operaejonalizacji teoretycznych zmiennych eksperymentalnych — sposób operacjonalizacji badacz zwykle musi obmyśleć na użytek konkretnej serii eksperymentów. Otwiera się tu przed nim rozległe pole możliwości — jest ono polem do popisu i wykazania własnej pomysłowości. „Można powiedzieć — piszą Aronson i Carlsmith — że sztuka eksperymentowania polega w pierwszym rzędzie na sprawności badacza w określaniu procedury, która jest najbardziej dokładną realizacją jego zmiennej teoretycznej oraz ma największy wpływ na badanych i budzi z ich strony najmniej podejrzeń”[33]. Sławni eksperymentatorzy, których nazwiska często będziemy cytować, tacy jak Asch, Hovlahd czy Milgram rozgłos swój zawdzięczają w ogromnym stopniu pomysłowości swych manipulacji.

Twórcza improwizacja w tej dziedzinie ma również ujemne strony. Jeśli eksperymentator wymyśla na użytek swych eksperymentów rozmaite operacjonalizacje zmiennych teoretycznych i wybiera spośród nich tę, która intuicyjnie wydaje mu się najbardziej trafną, to pojawia się możliwość, że wybrana realizacja zmiennej trafną nie będzie, a jej ewentualny różnicujący wpływ na zachowanie badanych jednostek będzie pochodził od tych jej aspektów, które nie są związane ze zmienną teoretyczną, od aspektów teoretycznie nieistotnych, przypadkowych. Niebezpieczeństwo to jest tym większe, im bardziej złożone są realizacje

zmiennych teoretycznych; realizacje takie obciążone są bowiem wielorakim znaczeniem. „Problem wielorakiego znaczenia jest — zdaniem kompetentnych eksperymentatorów i metodologów — jednym z najbardziej kłopotliwych problemów w eksperymentowaniu w psychologii społecznej. Przede wszystkim dlatego, że manipulacje, które wydają się badanym najbardziej sensownymi i wywierają na nich największy wpływ, zazwyczaj składają się ze złożonej wiązki bodźców i w konsekwencji — operacje możną zinterpretować na więcej niż jeden sposób”[34]. Powstaje pytanie, czy za zmianę zmiennej zależnej odpowiedzialny jest ten spośród wiązki bodźców, który koresponduje ze zmienną teoretyczną?

Twierdzącą odpowiedź na to pytanie rzadko się uzasadnia. Wydaje się, że przekonanie eksperymentatora o trafności manipulacji wynika zazwyczaj po prostu z jego wiedzy teoretycznej i jego własnego niepowtarzalnego doświadczenia badawczego. Ten ryzykowny rodzaj uzasadnień znany jest doskonale nie tylko eksperymentatorom, lecz także np. badaczom przekładającym zmienne teoretyczne na język pytań i odpowiedzi kwestionariuszowych. Badacze ci stosują do oceny trafności pytań tzw. metodę logiczną, która prawdę mówiąc, nie jest ani metodą, ani logiczną a polega po prostu na dociekaniu sensu pojęć teoretycznych i ich kwestionariuszowej operacjonalizacji.

Eksperymentatorom niepewnym teoretycznego znaczenia swych lub cudzych manipulacji doświadczeni badacze zalecają procedurę oczyszczania. Polega ona na wielokrotnym powtarzaniu eksperymentu z równoczesnym różnicowaniem tych aspektów bodźca, które są teoretycznie nieistotne. Różnicowanie takie można przeprowadzać nie tylko między różnymi eksperymentami, lecz także w obrębie jednego, badania. W eksperymentach nad perswazją można modyfikować np. temat przekazu z równoczesnym zachowaniem jego perswazyjnego charakteru albo rodzaj źródła z równoczesnym zachowaniem — powiedzmy — poziomu prestiżu źródła. Jeśli w tak zróżnicowanych doświadczeniach uzyska się te same wyniki,

[32] W dalszych rozważaniach w poważnym stopniu opieramy się na wnikliwym studium E. Aronsona i J.M. Carlsmitha, Experimentation in social psychology, w: G. Lindzey, E. Aronson (eds.), The handbook of social psychology, Reading, Mass, 1968, vol. 2. W r. 1976 ukazała się książka Methods of research in social psychology, pióra J. M. Carlsmitha, P. C. Ellswortha i E. Aronsona, będąca rozwinięciem idei cytowanego studium.

[33] Ibidem, s. 25.

[34] Ibidem, s. 13.

uprawdopodobni to wniosek, że różnice zachowania między badanymi wynikają z teoretycznie istotnych, a nie z przypadkowych aspektów manipulacji.

Maksymalna standaryzacja manipulacji eksperymentalnych nie jest więc tak cenna, jak się może wydawać. Istnieje jeszcze drugi ważny powód różnicowania manipulacji eksperymentalnych. W eksperymentach realizowanych w wielu zróżnicowanych środowiskach, różnicowanie manipulacji jest konieczne, jeśli ich sens teoretyczny ma być identyczny. Manipulacja powinna stawiać wszystkie badane jednostki w jednakowym teoretycznie opisanym położeniu. Nie wynika stąd jednak, że wszystkie jednostki powinny otrzymywać jednakowe bodźce. Standaryzacja na poziomie teoretycznym niejednokrotnie może być osiągnięta tylko przez brak standaryzacji na poziomie operacyjnym. Podobnie rzecz się ma z teoretyczną interpretacją wyników pomiaru. Zaobserwowane różnice w wynikach między różnymi środowiskami wcale nie muszą oznaczać środowiskowego zróżnicowania badanych zjawisk — mogą one wynikać po prostu ze środowiskowych stylów ekspresji.

Wnioski te znajdują szczególnie przekonujące świadectwa w eksperymentalnych międzykulturowych badaniach porównawczych. Przykładem takich badań są badania nad uwarunkowaniami odrzucania dewiantów, wykonane na początku lat 50-tych w krajach Europy Zachodniej[35]. Jedna z hipotez sprawdzanych w tych badaniach głosiła, że im większa atrakcyjność celu dla grupy, tym silniej grupa będzie odrzucać dewiantów blokujących osiąganie celu. Teoretyczną zmienną niezależną — atrakcyjność celu trzeba było zoperacjonalizować tak, by zapewnić teoretyczną standaryzację na poziomie międzykrajowym. Okazało się to zadaniem niełatwym.

Grupami eksperymentalnymi były kluby modelarstwa lotniczego, skupiające kilkunastoletnich chłopców. Jednakże utworzenie takich klubów w Republice Federalnej Niemiec napotkało na nieoczekiwane trudności. W owych latach w Republice Federalnej obowiązywał jeszcze status okupacyjny a ledwie kilka miesięcy przed rozpoczęciem eksperymentu zniesiono zakaz wszelkiej działalności związanej z lotnictwem. W tej sytuacji próba rekrutacji chłopców do klubów lotniczych wywołała zaniepokojenie władz szkolnych, rodziców a nawet dochodzenie policyjne. Mimo oficjalnych wyjaśnień, reakcje te nie ustały i poważnie zakłócały przebieg eksperymentu. Trudno było wymyślić bardziej nieodpowiedni kontekst eksperymentalny w Republice Federalnej — przyznali to sami eksperymentatorzy.

W eksperymentach tych celem o niskiej atrakcyjności były bilety na nudne filmy dokumentalne, uzyskiwane w nagrodę za pracę modelarską. Okazało się jednak, że w Belgii cel ten ceniony był bardzo wysoko, gdyż dla znacznej części chłopców, chodzących do szkół parafialnych, pójście do kina było rzadkim przywilejem. Autorzy tego eksperymentu żałowali, że nie zastosowali w Belgii innej operacjonalizacji niż w pozostałych krajach, np. nudnej książki i stwierdzili, że w pewnych przypadkach standaryzację międzykrajową najlepiej osiągnąć kosztem uniformizacji sytuacji bodźcowych.

Efektywna kontrola hipotezy teoretycznej wymaga nie tylko znalezienia trafnych operacyjnych odpowiedników teoretycznej zmiennej niezależnej, lecz także skutecznej realizacji tych zdarzeń. Eksperymentator powinien się upewnić, czy przy pomocy swych manipulacji wpłynął na poziom zmiennej niezależnej, czy udało mu się stworzyć warunki, w których hipoteza teoretyczna będzie mogła być' sprawdzona. Może się bowiem okazać, że to zdarzenie, trafnie wskazane jako operacjonalizacja pewnego poziomu zmiennej teoretycznej, nie zostanie faktycznie zrealizowane, a wówczas wniosek o fałszywości sprawdzanej hipotezy sam będzie fałszywy.

Eksperyment wykonany przez Greenberga dobrze ilustruje to niebezpieczeństwo[36]. W psychologii społecznej znane są badania Schachtera nad społecznymi

[35] Zob.: Cross-national research: a case study, „Journal of Social Issues" 1954, vol. 7, nr 4; instruktywne są zwłaszcza prace Schachtera oraz Rommetveita i Israela.

[36] M. S. Greenberg, Role playing — an alternative to deception?, w: A. G. Miller (ed.), The social psychology of psychological research, New York 1972.

uwarunkowaniami potrzeby afiliacji, rozumianej jako dążenie jednostki do przebywania z innymi ludźmi, gdy znajduje się ona w sytuacji niejasnej. Schachter wykonał eksperyment, w którym manipulował poziomem lęku (sytuacje wywołujące lęk są sytuacjami niejasnymi) przez informowanie badanych, że będą poddani silnym lub słabym wstrząsom elektrycznym, następnie zapowiadał odłożenie eksperymentu na pewien czas i pytał, czy badani chcą czekać na wstrząsy w samotności czy w towarzystwie innych — w ten sposób mierzył siłę potrzeby afiliacji. Schachter wykazał, że w warunkach silnego lęku potrzeba afiliacji ujawnia się mocniej u jedynaków i pierworodnych niż u dzieci urodzonych w dalszej kolejności i że wśród jedynaków i pierworodnych wyższy poziom lęku powoduje silniejsze dążenie do afiliacji. Eksperyment Greenberga tym różnił się od doświadczenia Schachtera, że eksperymentator zamiast straszyć badanych serią szoków elektrycznych, poprosił ich, by postawili siebie w sytuacji oczekiwania na serię wstrząsów i odpowiedzieli, czy chcieliby wówczas czekać w samotności czy razem z innymi — był to eksperyment posługujący się techniką odgrywania ról. Okazało się, że porządek urodzenia nie różnicował odpowiedzi, gdy drugą ze zmiennych niezależnych był poziom lęku określany przez eksperymentatora. Natomiast różnice wystąpiły, gdy za zmienną wzięto poziom lęku percepowany przez badanych. Ciekawe, że eksperyment Greenberga dał takie same wyniki jak badania Schachtera. Gdyby jednak Greenberg nie zapytał badanych, co czuliby oczekując na wstrząsy, to uzyskałby wyniki przeczące hipotezie.

Istnieje wiele metod pozwalających sprawdzić skuteczność manipulacji.

Pierwszą z nich jest pomiar zmiennej niezależnej zaraz po wprowadzeniu bodźców. Tak np. w eksperymencie Morse'a i Reimera nad wpływem zmian w stopniu uczestnictwa pracowników w podejmowaniu decyzji organizacyjnych na wydajność pracy i zadowolenie z pracy[37], sprawdzono przez obserwację zachowania pracowników i badanie kwestionariuszowe, czy formalne zmiany w strukturze władzy spowodowały faktyczne zmiany zachowania i czy zostały dostrzeżone. Pomiary tego rodzaju nie zawsze są możliwe i nie zawsze są wskazane, gdyż mogą ujawnić fakt lub cel eksperymentu. I tak w omawianym eksperymencie okazało się, że wydajność pracy wzrosła zarówno w działach, w których autonomię pracowników zwiększono jak i tam, gdzie ją zmiejszono. Eksperymentatorzy podejrzewali, że przyczyną tego mogły być pomiary eksperymentalne, odczytane przez pracowników jako wzrost zainteresowania nimi ze strony dyrekcji; świadomość bycia przedmiotem zainteresowania mogła motywować do wydajniejszej pracy. Był to przykład zakłócenia eksperymentalnego zwanego efektem Hawthorne.

Druga z metod polega na pomiarze zmiennych interweniujących w procesie wpływu zmiennej niezależnej na zależną. W hipotetycznym eksperymencie nad wpływem frustracji na przemieszczenie agresji, można byłoby starać się sprawdzić, czy wystąpiła agresja — zmienna pośrednicząca między frustracją a przemieszczaniem agresji. Gdyby wystąpiła, to pewność, że manipulacja eksperymentalna była frustrującą podniosłaby się znacznie.

Trzecią metodą jest wywiad po zakończeniu eksperymentu. Może on dostarczyć istotnych informacji pozwalających ocenić skuteczność zastosowanych manipulacji. Eksperymentator może się dowiedzieć, czy jego operacje zostały potraktowane przez badanych na serio i czy odniosły swój skutek. Dobrym przykładem są tu badania Milgrama nad gotowością do zadawania cierpienia wbrew etyce, lecz zgodnie z rozkazem. O badaniach tych będziemy mówić obszerniej jeszcze wiele razy. W tym miejscu wskażemy tylko, że ich autor przekonał się w czasie wywiadu poeksperymentalnego, że jego polecenia zostały przyjęte poważnie i badani działali w przekonaniu, że cierpienia zadają[38].

[37] N. C. Morse, E. Reimer, The experimental change of a major organizational variable, w: W/M. Evan (ed.), Organizational experiments, New York 1971.

[38] Zob. S. Milgram, Some conditions of obedience and disobedience to authority oraz Interpreting obedience; error and evidence, w: The individual in a social world, New York 1978.

Czwartą z metod mogłyby być specjalne eksperymenty, których celem byłoby wyłącznie sprawdzenie skuteczności manipulacji. Wiedza z nich uzyskana byłaby cenną dla prowadzenia pozostałych z planowanej serii badań.

Zauważmy na koniec, że czasem zgodność uzyskanych wyników z dobrze uzasadnionymi teoriami przekonuje eksperymentatora, że zmienne teoretyczne zoperacjonalizował on trafnie i operacjonalizacje wprowadził skutecznie.

5. ZAKŁÓCENIA BODŹCA EKSPERYMENTALNEGO

Zapewnienie eksperymentowi trafności wewnętrznej nie jest zadaniem łatwym. Na działanie bodźca nakłada się bowiem wpływ innych czynników — czynniki te zanieczyszczają bodziec, są źródłami jego kontami- nacji. Mogą one spowodować fałszywą ocenę wpływu bodźca na zmienną zależną. Jeśli działają w kierunku zgodnym z bodźcem, to mogą wieść do przecenienia tego wpływu — do przypisania bodźcowi zmian nie przez niego a przez siebie wywołanych. Jeśli zaś działają w kierunku przeciwnym, mogą doprowadzić do niedocenienia siły bodźca — niezauważenia zmian przez siebie stłumionych. Częstszą jest pierwsza z tych sytuacji i nią głównie będziemy się zajmować.

Przedstawimy teraz listę zakłóceń bodźców eksperymentalnych. Listę tę sporządził Campbell[39] W literaturze metodologicznej funkcjonuje ona jako klasyczna.

Powiedzmy, że w celu sprawdzenia hipotezy teoretycznej o związku przyczynowym między dwiema zmiennymi — niezależną a zależną, zaprojektowaliśmy i wykonaliśmy pewien eksperyment. Eksperyment ów polegał na umyślnej zmianie wartości zmiennej niezależnej, określonej na pewnym Zbiorze jednostek w celu sprawdzenia, czy wartość zmiennej zależnej zmieni się tak, jak głosi hipoteza. Zakładamy przy tym, że przekład zmiennych teoretycznych, o których mówi hipoteza, na manipulacje eksperymentalne i operacje pomiarowe został dokonany trafnie.

Eksperyment ten możemy zaprojektować na /wiele różnych sposobów. W tym miejscu jednakże wystarczy wymienić tylko dwa z nich. Możemy więc dobrać dwie grupy porównawcze, jedną z nich — eksperymentalną — poddać manipulacji, a drugą — kontrolną — nie i, po pewnym czasie zmierzyć i porównać wartości zmiennej zależnej w obu grupach. Możemy też zadowolić się tylko jedną grupą: wówczas musimy zmierzyć wartość zmiennej zależnej, wprowadzić bodziec, po pewnym czasie zmierzyć ponownie zmienną zależną i porównać uzyskane wyniki pomiarów. Z otrzymanych w tych eksperymentach różnic będziemy wnosić o wpływie manipulacji na zmianę wartości zmiennej zależnej i — dalej — o wartości logicznej sprawdzanej hipotezy.

Od stwierdzenia takich różnic do decyzji o wartości hipotezy droga jest jednak daleka.

Uzyskane różnice mogą być dziełem przypadku tj. niekontrolowanych zdarzeń losowych; ten rodzaj zakłóceń bodźca figuruje na liście Campbella jako niestabilność.

Zakłócenia przypadkowe obejmują wiele czynników. Między grupami porównawczymi istnieją pewne różnice w początkowych wartościach zmiennej zależnej i zmiennych istotnych, nawet jeśli dobrano je losowo. Grupy te nie są wewnętrznie jednorodne, co ma znaczenie wtedy, gdy metodą analizy wyników jest ana-

[39] Zob. D. T. Campbell. Factors relevant to the validity of experiments in social settings (1957) w: N.K. Denzin (ed.), Sociological methods, London 1970; D. T. Campbell, J. C. Stanley, Experimental and quasi-experimental designs for research, Chicago 1973; D. T. Campbell, Reforms as experiments, w: E. L. Struening, M. Guttentag (eds.), Handbook of evaluation research, Beverly Hills 1975; Th. D. Cook, D. T. Campbell, The design and conduct of quasi-experiment and true experiments in field settings, w: M. D. Dunnette (ed.), Handbook of industrial and organizational psychology, Chicago 1976. Wymienione prace pozwalają prześledzić rozwój tej listy.

liza wariancji. Warunki, w jakich działają obie grupy i warunki, w jakich działają poszczególne jednostki w ich obrębie nie są identyczne. Mogą istnieć drobne, przypadkowe różnice w sposobie wprowadzania lub percepcji bodźca u poszczególnych jednostek. Narzędzia pomiaru nie są idealnie rzetelne — znaczy to m.in., że są one wrażliwe w pewnym stopniu na przypadkowe okoliczności pomiarów. Wszystkie te zakłócenia tłumaczą pewną część zmienności uzyskanych wyników — nazywa się ją błędem eksperymentalnym.

Niejednokrotnie wystarczy rzut oka na uzyskane różnice, by nabrać pewności, że nie mogły ich wywołać czynniki losowe. Pewniejszą i bardziej uniwersalną metodą dochodzenia do takich wniosków jest jednak stosowanie statystycznych testów istotności. Testy istotności pozwalają stwierdzić, czy uzyskane różnice przekraczają wielkości, których należało oczekiwać na zasadzie przypadku i tym samym pozwalają odrzucić z małym ryzykiem błędu hipotezę o losowym charakterze tych różnic.

Z samego stwierdzenia istotności różnic nie wynikają jednak żadne wnioski co do czynników, które Spowodowały te różnice, w szczególności nie wynika nic, co by dotyczyło wpływu bodźca (zmiennej niezależnej) na zmienną zależną. Test istotności informuje tylko, że między zmienną niezależną a zależną istnieje korelacja. Taka informacja jest konieczna do dowiedzenia związku przyczynowego, lecz nie jest dlań wystarczająca. Niezbędna jest również kontrola innych zmiennych systematycznie powiązanych ze zmienną niezależną. Jeśli bierze się machinalnie pozytywny wynik testu istotności za dowód wpływu bodźca, to widocznie przyjmuje się założenie, że eksperyment jest tak dobrze kontrolowany, że między grupami istnieją tylko różnice przypadkowe i różnice bodźca.

Przypadek nie jest jednak jedynym źródłem kontaminacji bodźca — ani jedynym możliwym, ani jedynym spotykanym.

Drugim rodzajem zakłóceń są zmiany zewnętrzne; Campbell nazywa je historią. Nazwa ta oznacza wszystkie wydarzenia zachodzące w środowisku podczas eksperymentu, które mogą wpłynąć na wystąpienie lub wielkość zmiany eksperymentalnej. W badaniach rynkowych, w których oceniano by skuteczność reklamy pewnego towaru, pojawienie się na rynku towaru konkurencyjnego mogłoby zmniejszyć zainteresowanie towarem reklamowanym tak, że reklama zostałaby uznana nie tylko za nieskuteczną, lecz wręcz za szkodliwą. W badaniach skuteczności perswazji, prowadzonych w środowisku naturalnym, równoczesne działanie innych treści, propagowanych przez źródła nieeksperymentalne, mogłoby podnieść lub obniżyć atrakcyjność propagowanych poglądów. W eksperymentalnych badaniach efektywności nowych sposobów zarządzania w przemyśle, zmiana warunków kooperacji czy zbytu produkowanych towarów mogłaby wpłynąć na przecenienie lub niedocenienie wartości innowacji.

Bez wątpienia, w wielu wypadkach zakłócenia należące do tej kategorii wolno zlekceważyć, zwłaszcza, gdy prowadzi się krótkotrwałe badania psychologiczne w laboratorium. Ale w środowisku naturalnym, w badaniach długotrwałych, niebezpieczeństwo tego rodzaju kontaminacji jest bardzo poważne.

Trzecia grupa czynników zakłócających obejmuje samorzutne zmiany badanych układów, nie wymagające zewnętrznych bodźców w czasie trwania eksperymentu. Campbell nazywa je dojrzewaniem, gdyż dojrzewanie rozumiane jako proces wzrostu osobniczego, stosunkowo niezależny w swym przebiegu od środowiska, stanowi dobry przykład takich zmian, aczkolwiek w wielu przypadkach właściwszy byłby termin „trend”. Zmiany dyskutowanego rodzaju występują pod dwiema postaciami:

Po pierwsze, zjawiska i systemy badane przez socjologów odznaczają się tym typem autodynamiki, który polega na doświadczaniu zmian w miarę upływu czasu, bez wpływu wydarzeń zewnętrznych. „Czas” nie występuje tu w roli samodzielnej zmiennej, lecz w roli wygodnego kryptonimu dla procesów biologicznych, psychologicznych i socjologicznych, które rozwijają się w pewnym rytmie. Starzenie się. dojrzewanie, biurokratyzacja struktur organizacyjnych, rozmaite fluktuacje sezonowe są przykładami takich procesów.

Po drugie, układy społeczne przechodzą w czasie trwania eksperymentu zmiany spowodowane uprzednim działaniem na nie różnych okoliczności. W czasie trwania eksperymentu z zakresu nowych technik zarządzania, może następować systematyczna poprawa wskaźników sprawności instytucji, nie mająca niczego wspólnego ze zmianami wprowadzonymi eksperymentalnie. Może on natomiast wynikać z podjętych wcześniej trafnych decyzji lub też może być fragmentem długotrwałego i ogólnego trendu.

Czynniki te, nie rozpoznane i nie skontrolowane w porę, mogą stać się powodem błędnej interpretacji wyników eksperymentu, przypisującej bodźcom wpływy zmian samorzutnych np. nowej technice dydaktycznej — efekty dojrzewania umysłowego, a reklamie — działanie mody. I znowu, jak w przypadku historii, w wielu eksperymentach niebezpieczeństwo to można pominąć, lecz w wielu innych trzeba się z nim uporać.

Czwartym czynnikiem zakłócającym jest sam pretest. Różnice między pomiarem wstępnym a końcowym można wyjaśnić nie wpływem bodźca, lecz alternatywnie — wpływem pierwszego z pomiarów. Sam akt pomiaru, jeśli tylko badani są go świadomi lub też sam wynik pomiaru, jeśli badani go znają, może wpływać na wyniki pomiarów następnych. Rozmaite bywają mechanizmy pośredniczące w tym wpływie.

a) Niekiedy doświadczenie nabyte przy rozwiązywaniu zadań testowych, oswojenie się z narzędziem pomiaru, wpływa dodatnio na wyniki następnych pomiarów przy pomocy tego samego testu lub jego odmiany. Tak np. w pomiarach inteligencji testem IQ doświadczenie zyskane w pierwszym pomiarze powoduje osiągnięcie w drugim pomiarze wyniku lepszego o kilka punktów. Podobny efekt zaobserwowano w wypadku niektórych testów osiągnięć i przystosowania się: badani pierwszy raz uzyskiwali wyniki nieco gorsze niż badani drugi raz[40]. Łatwo sobie wyobrazić trudność, jaką sprawia ten efekt w badaniach np. nad uczeniem się, skoro i on sam jest niczym innym jak uczeniem się — nabywaniem pewnych umiejętności pod wpływem poprzednich doświadczeń. Trudno wówczas powiedzieć, jaka część różnicy między pomiarami uczonych reakcji wzięła się z treningu, a jaka z pretestu samego.

b) Innym razem sam fakt bycia badanym zmienia ludzi pod tymi względami, które są przedmiotem badania. Przede wszystkim aktywizuje ich poznawczo: zwraca ich uwagę na pewne sprawy, skłania do szukania nowych o nich informacji, zachęca do myślenia o rzeczach, o których dotąd być może nawet nie myśleli, ukazuje nowe możliwości patrzenia na jakieś zagadnienia. Wiadomo zaś, że stopień znajomości pewnych spraw ma wpływ na postawy względem nich zajmowane. Na tym m. in. polega efekt panelowy. Tak np. w badaniach panelowych kampanii wyborczej w r. 1961 Gostkowski wykazał, że samo tylko przeprowadzenie wywiadu spowodowało przyrost wiedzy i zmianę postaw niezależnie od wpływu kampanii propagandowej[41]

c) Z jeszcze inną sytuacją mamy do czynienia wtedy, gdy badani znają wyniki pretestu. Mogą one stanowić dla nich bardzo ważną informację, uświadamiającą im stopień rozbieżności między rzeczywistym stanem rzeczy a ich standardami i skłaniającą ich do działań, zmierzających do ich uzgodnienia. Zdanie sobie sprawy z niskiego poziomu wiedzy może zachęcić uczniów do intensywniejszej pracy i spowodować jego podniesienie się niezależnie od działania eksperymentalnej metody dydaktycznej. Rzecz jasna, niektóre tylko cechy są podatne na tego rodzaju oddziaływania. Wynik pretestu jest tu nie tylko biernym zapisem stanu rzeczy, lecz także bodźcem do zmiany tego stanu. Omawiany mechanizm istotny jest zwłaszcza w eksperymentalnych próbach innowacji społecznych.

Zauważmy, że wpływ pretestu na posttest polega raz na bezpośrednim wpływie aktu pomiaru na reakcje testowe w poststeście (a) a innym razem na wpływie aktu

[40] Zob.: Campbell, Stanley, op. cit., s. 9.

[41] Z. Gostkowski, Analiza „efektu panelowego" w badaniach wyborczych w Łodzi w r. 1961, w: Z. Gostkowski (red.), Analizy i próby technik badawczych w socjologii, t. 1., Wrocław 1966.

pomiaru lub jego wyniku na zmienną zależną, co posttest tylko rejestruje (b,c). Tylko w drugim przypadku istnieje faktyczna zmiana, w pierwszym istnieje tylko jej pozór. Zanotujmy jeszcze, że dyskutowane wpływy nie mają miejsca wtedy, gdy pomiary są tak dyskretne, że badani nie zdają sobie z nich sprawy, a jeśli już, to nie znają ich wyniku.

Wyłączyliśmy z rozważań ten rodzaj wpływu pretestu na zmienną zależną, który polega na warunkowaniu wpływu bodźca; o interakcji pretestu z bodźcem powiemy w odpowiednim miejscu.

Regresja stanowi kolejny, piąty rodzaj zakłóceń bodźca. Efekt regresji występuje wówczas, gdy względem, który decyduje o tym, że pewien obiekt zostaje poddany interwencji eksperymentalnej, jest posiadanie przezeń w chwili badania skrajnej wartości zmiennej zależnej, stwierdzone w preteście. Dzieje się tak wówczas, gdy bodźcowi poddaje się tę część zbiorowości lub przedmiotem analizy czyni się tę część grupy eksperymentalnej, która w preteście uzyskała krańcowe wyniki np. zdecydowanych przeciwników jakiegoś poglądu, uczniów dwójkowych. Dzieje się tak również wtedy, gdy interwencja eksperymentalna w jakiś zmieniający się w czasie układ następuje w chwili, gdy wartość charakteryzującej go zmiennej szczególnie mocno odchyla się od ogólnego trendu, np. gdy fabryka przechodzi nagły ostry kryzys wydajności pracy. W pierwszym wypadku *wybiera* się krańcowe obiekty ze zbioru obiektów a w drugim krańcowe stany spośród zbioru stanów tego samego obiektu. Należy wówczas oczekiwać, że wyniki posttestu będą mniej krańcowe — przesuną się ku średniej bez jakiegokolwiek wpływu bodźca. Grupy krańcowe przesuną się ku średniej populacji, z której je wzięto, a krzywa, oddalona od trendu, przybliży się doń. Na tym polega efekt regresji.

Efekt regresji wyjaśnia się zwykle w języku błędu pomiaru. „Im bardziej wynik odchyla się od normy, tym większy błąd pomiaru prawdopodobnie zawiera — piszą Campbell i Stanley. Tak więc, typowy osobnik ze skrajnie wysokim wynikiem miał w pewnym sensie niezwykłe szczęście (duży błąd pozytywny), a osobnik z wynikiem skrajnie niskim miał niezwykłego pecha (duży błąd negatywny). Los jest jednak kapryśny, więc oczekujemy, że w postteście pierwsi opadną nieco ku średniej, a drudzy poprawią swe względne położenie. (Ta sama logika obowiązuje, gdy zaczyna się od wyników posttestu i cofa się do protestu)”[42]. Dyskutowany efekt prowadzi do tego, że zmienia się skład grup krańcowych, lecz rozkład zmiennej w całej zbiorowości wcale zmieniać się nie musi. Dodatkowy aspekt sprawy polega na tym, że jeśli grupy krańcowe zajmują pozycje skrajne na skali zamkniętej w sposób konwencjonalny od góry i od dołu, to skala ta odnotuje tylko przesunięcia do średniej. Uczniowie dwójkowi mogą tylko poprawić swoje wyniki, a piątkowi tylko pogorszyć.

Regresja w tym ujęciu stanowi jedynie tautologiczną implikację niedoskonałej rzetelności narzędzia pomiaru, wyrażonej mniejszą od jedności korelacją wyników dwóch pomiarów tego samego obiektu. Im korelacja niższa i na mocy definicji niższa rzetelność pomiaru, tym większa regresja. Do pojawienia się pewnej różnicy między posttestem i pretestem w eksperymencie na grupie skrajnej nie jest więc wcale konieczny bodziec.

Ale regresję można niekiedy wyjaśnić w kategoriach zmian rzeczywistych. Tak czyni Coleman[43], a także Kessler[44]. W pewnych systemach społecznych zmiana wartości jakiejś zmiennej nie zależy od wartości innych zmiennych, lecz od początkowej wartości samej tej zmiennej. Szczególnym przypadkiem takich systemów są układy samoregulujące się, homeostatyczne. W układach takich splot ujemnych sprzężeń zwrotnych sprawia, że po znacznych odchyleniach zmiennych od stanu równowagi, następuje ich samoczynny powrót do tego stanu. Powrót ten to właśnie regresja. Mechanizm ten dobrze wyjaśnia zjawisko przybliżania się do

[42] Campbell, Stanley, op. cit., s. 11.

[43] J. S. Coleman, Mathematical study of change, w: H. M. Blalock Jr., A. B. Blalock (eds.), Methodology in social research, New York 1968.

[44] R. C. Kessler, The use of change scores as criteria in longitudinal survey research, „Quality and Quantity” 1977, vol. 11, nr 1.

trendu ostrych odeń odchyleń, obserwowane na wykresach obrazujących zmiany rozmaitych zjawisk w czasie. Bezkrytycznemu eksperymentatorowi interwencja w działanie systemu samoregulującego się wykonana w chwili zakłócenia Jego równowagi zawsze więc wydałaby się skuteczna — tymczasem zmiana nastąpiłaby samoczynnie nawet wtedy, gdyby manipulacja nie miała na system żadnego wpływu. Do takiego ujęcia regresji i tłumaczenia nią zmiany konieczne jest uprzednie dowiedzenie, że badany system jako całość posiada własność samoregulacji. Ciekawe jest też wiedzieć, w jaki to sposób systemowy mechanizm równowagi reguluje zachowania elementów systemu.

Szóstym rodzajem zakłóceń jest zmiana zasad i narzędzi pomiaru w okresie między pretestem i posttestem. Może wówczas zdarzyć się nawet taka sytuacja, w której za różnicą pomiarów nie będzie się kryć jakakolwiek zmiana w wartościach zmiennej zależnej, lecz tylko zmiana w samych metodach pomiaru.

Zmiany tego rodzaju nie są w socjologii czymś wyjątkowym. W nauce tej narzędziem pomiaru czy lepiej — jego elementem jest sam człowiek, występujący w roli obserwatora, ankietera, sędziego, kodera, urzędnika spisowego. Dokonujące się z upływem czasu zmiany jego wrażliwości na pewne fakty, sprawności ich zapisu, jego kryteriów interpretacji zjawisk, mogą stworzyć pozór zmiany rzeczywistości. Kilka przykładów pokaże, że badacz nie jest tym samym narzędziem w różnych punktach czasowych badania.

Goode i Hatt notują, że obserwator „w miarę przyswajania sobie «poprawnych» sposobów zachowania się, zaczyna uważać je za tak oczywiste, że wydają mu się całkowicie naturalne, wskutek czego nie zwróci uwagi na wiele szczegółów. Będą dla niego tak powszednie, że aż nie warte uwagi”[45]. Brzozowska-Ejme i Kawka wykonały badania nad ankieterami i wykazały, jak wrażliwość na zachowanie respondentów i staranność zapisu wywiadu najpierw rośnie a potem, po przekroczeniu pewnego punktu respondenci „stają się” coraz bardziej do siebie podobni, a zapis staje się coraz mniej dokładny[46]. Źródłem różnic może być także zmiana w zasadach działalności i sprawności pozanaukowych instytucji zbierających dane. Krzywa obrazująca rzekomy spadek przestępczości może faktycznie wyrażać tylko zaostrzenie kryteriów przestępstwa i (lub) spadek wykrywalności przestępstw — „ciemna liczba” pozostaje bez zmian.

Dopiero, gdy narzędzia pomiaru będą niezmienne, wolno będzie z różnic ich odczytów wnosić o zmianach rzeczywistości. Stąd dylemat stojący przed badaczami zmiany społecznej: czy lepiej mieć narzędzia gorsze, lecz łatwiej porównywalne czy lepsze lecz trudniej porównywalne, rozstrzygać trzeba raczej na korzyść pierwszego wyjścia. Nie należy udoskonalać narzędzi w trakcie eksperymentu, jeśli miałoby to obniżyć porównywalność wyników.

Selektywny dobór grup w eksperymencie jest siódmym rodzajem zakłóceń bodźca.

W eksperymentach posługujących się więcej niż jedną grupą czyli niemal we wszystkich eksperymentach, grupy muszą być ujednolicone ze względu na wartości przynajmniej zmiennej zależnej i zmiennych istotnych. W przeciwnym razie, różnice między grupami stwierdzone w postteście będą mogły być wyjaśnione nie tylko różnicami bodźców, lecz również — konkurencyjnie — różnicami początkowymi. Siła wnioskowania przyczynowego z danych eksperymentalnych jest bardzo duża m. in. i z tego powodu, że kontakt jednostek z bodźcem i jego brak są uniezależnione od własności jednostek. Uniezależnienie to ’najlepiej osiąga się przez losowy dobór jednostek do grup w eksperymencie.

Nie ma tej niezależności wtedy, gdy grupę eksperymentalną tworzy się *ex post* z jednostek, które już się z bodźcem zetknęły a kontrolną z jednostek, które się z bodźcem nie zetknęły. Nie ma tej niezależności również wówczas, gdy daną zbio-

[45] W. J. Goode, P.K. Hatt, Obserwacja, w: S. Nowak (red.), Metody badań socjologicznych, Warszawa 1965, s. 50,

[46] K. Brzozowska - Ejme, Z. Kawka, Analiza błędów pracy ankietera w wywiadzie kwestionariuszowym na podstawie zapisu magnetofonowego, w: Z. Gostkowski i J. Lutyński (red.) Analizy i próby technik badawczych w socjologii, Wrocław 1970, t. 3.

rowość dzieli się na grupę eksperymentalną i kontrolną na podstawie preferencji jednostek. W obu wypadkach porównanie między grupami jest nadzwyczaj ryzykowną miarą wpływu bodźca. W sytuacjach, w których dobór samorzutny decyduje o kontakcie z bodźcem, pojawiają się pytania: Czy o kontakcie tym decydują okoliczności przypadkowe, nieistotne dla wpływu bodźca czy też przeciwnie — decydują uprzednie wartości zmiennej zależnej i czynniki dla wpływu bodźca istotne?

Wiele wskazuje na to, że zwykle dobór ten obciążony jest systematycznym błędem autoselekcji. W grupie eksperymentalnej znajdują się jednostki o innych wartościach zmiennej zależnej i zmiennych istotnych niż w grupie kontrolnej. Gdyby np. w terenowym eksperymencie nad skutecznością oddziaływania reklamy prasowej porównano jednostki, które zetknęły się z przekazem z tymi, które się z nim nie zetknęły, to najprawdopodobniej uzyskano by błędne wnioski, gdyż sam kontakt byłby różnicowany przez używanie reklamowanego towaru oraz takie istotne dla wpływu zmienne, jak wiek czy wykształcenie. W pewnym eksperymencie badano wpływ przebywania studentów przez dłuższy czas w dzielnicy murzyńskiej na redukcję uprzedzeń etnicznych. Grupę eksperymentalną stworzono z osób, które chciały do niej wejść[47]. Bez wątpienia ich uprzednie postawy etniczne nie były obojętne ani dla tej decyzji, ani dla wpływu bodźca. Do tego źródła kontaminacji będziemy mieli jeszcze okazję wrócić.

Ósmym, ostatnim rodzajem zakłóceń jest ubywanie elementów („śmiertelność") grup biorących udział w eksperymencie. Ubytek ten powodowany jest ruchliwością jednostek lub odmową dalszej współpracy z badaczem.

„Śmiertelność" ta rzadko tylko ma charakter losowy. Jeśli zmienna, ze względu na którą „śmiertelność" jest selektywna, związana jest ze zmienną zależną lufo też wręcz jest ową zmienną zależną, to bez jakiegokolwiek udziału bodźca powstaje sztuczne wrażenie zmiany: średnia wartość zmiennej zależnej ustalona w poststeście różni się od wartości wyliczonej w preteście. Jeśli np. z klasy szkolnej, w której sprawdzono eksperymentalnie wpływ nowej metody dydaktycznej na umiejętności uczniów, odpadnie w ciągu roku kilku uczniów słabych, to średni poziom umiejętności w klasie podniesie się, nawet jeśli nowa metoda nie różni się skutecznością od metod dotychczasowych.

Selektywne ubytki grupy eksperymentalnej można kontrolować przy pomocy grupy kontrolnej, jeśli w obu grupach dokonuje się ona ze względu na te same zmienne.

W celu określenia faktycznego wpływu bodźca trzeba więc upewnić się wpierw co do kierunku i wielkości ubytków z grup biorących udział w eksperymencie. Aby przyjąć hipotezę sprawdzaną w eksperymencie bez grupy kontrolnej należy odrzucić hipotezę konkurencyjną głoszącą, że różnice między pretestem a posttestem powstały w wyniku nielosowych zmian w składzie grupy, ażeby przyjąć tezę sprawdzaną w eksperymencie z grupą kontrolną należy odrzucić hipotezę głoszącą, że różnice między posttestami zostały spowodowane tym, że z każdej z grup jednostki wypadały wedle innych zasad.

Efekt dyskutowany nie występuje — rzecz jasna — w krótkotrwałych badaniach laboratoryjnych, w których eksperymentator kontroluje skład badanych grup. Występuje on natomiast w badaniach terenowych, prowadzonych na zbiorowościach naturalnych.

W ten sposób zamknęliśmy listę czynników konta- minujących wpływ bodźca na zmianę zależną. Lista jako całość stanowi jedynie wykaz zakłóceń możliwych, gdyż w konkretnych eksperymentach występują tylko niektóre z nich, a może się zdarzyć, że oprócz zakłóceń przypadkowych nie pojawią się żadne inne.

Działanie czynników ubocznych sprawia, że istotne różnice między wynikami pretestu a posttestu oraz istotne różnice między wynikami posttestów dadzą się w zasadzie wyjaśnić na wiele sposobów. Rzeczywiście różnice owe mogą wynikać z wpływu bodźca, tak jak głosi sprawdzana hipoteza. Jednakże różnica między pomiarem wstępnym a końcowym może być spowodowana wpływem zmian ze-

[47] Zob. E. Greenwood, Experimental sociology, New York 1945, s. 97 - 98.

wnętrznych, może być skutkiem wzrostu samorzutnego lub fragmentem trendu rozwojowego, może wynikać z wpływu samego pretestu, a jeśli grupa wybrana była ze względu na krańcową wartość zmiennej zależnej, różnica w wynikach kolejnych pomiarów może być spowodowana efektem regresji; różnica ta może też być wywołana selektywnym ubytkiem jednostek z grupy eksperymentalnej, wreszcie, może się zdarzyć, że za różnicami uzyskanych wyników nie kryje się żadna zmiana rzeczywista, lecz tylko zmiana zasad i narzędzi pomiaru. Różnice między wynikami posttestów w grupie eksperymentalnej i kontrolnej mogą być spowodowane różnicami w zasadach doboru jednostek do tych grup lub wypadania z nich. Każdy z tych kontaminujących bodziec czynników może więc być źródłem hipotez konkurencyjnych względem hipotezy sprawdzanej.

Po to, by z zaobserwowanych różnic wnioskować o wartości logicznej hipotezy sprawdzanej trzeba uprzednio wyeliminować możliwość owych wyjaśnień konkurencyjnych i zlikwidować dwuznaczność uzyskanych wyników. Wówczas wynik eksperymentu będzie jednoznaczny a eksperyment wewnętrznie trafny. Wykonanie tego zadania jest celem kontroli eksperymentalnej.

Trafność wewnętrzna nie jest jedynym kryterium wartości eksperymentu. Drugim ważnym kryterium jest trafność zewnętrzna, czyli możliwość uogólniania uzyskanych wyników. Nie wystarczy stwierdzić, nawet z niezbitą pewnością, że w pewnym eksperymencie bodziec wywołał takie to a takie skutki. Trzeba nadto znać zakres ogólności ustalonych związków oraz umieć projektować i prowadzić eksperymenty tak, by zakres ten był możliwie szeroki. Wchodzą tu w grę rozmaite wymiary uogólniania. Najważniejsze z nich to możliwość uogólnień ustalonych związków na inne zbiorowości niż zbiorowość, na której przeprowadzono eksperyment, na inne zewnętrzne warunki społeczne niż te, w których eksperyment się odbył i na inne sytuacje niż sytuacja eksperymentalna. Zależność, której ogólność chce się ocenić, została ustalona na szczególnej zbiorowości, w szczególnych warunkach zewnętrznych i w szczególnej sytuacji, jaką jest eksperyment. Każda z tych okoliczności może — nie wpływając na zmienną zależną — warunkować wpływ bodźca na tę zmienną.

Pasjonujące zagadnienie trafności zewnętrznej eksperymentu przedyskutujemy obszernie w odrębnym rozdziale. W tym miejscu postawiliśmy je tylko, gdyż przystępujemy do omawiania rozmaitych schematów eksperymentalnych, a schematy te różnią się często braniem i niebraniem pod uwagę pewnych czynników istotnych dla możliwości uogólniania stwierdzonych w eksperymentach zależności.

Powiedzieć jednak trzeba, że tym, co przede wszystkim różni od siebie poszczególne schematy eksperymentalne jest liczba bodźców, których wpływ można badać równocześnie oraz możliwości kontroli czynników zakłócających wpływ bodźców. Możliwości te zwiększa się właśnie przez obmyślanie coraz to bardziej pomysłowych *schematów* eksperymentalnych. Ogólność wyników trzeba zapewniać raczej inną drogą, choć i tej drogi — jak zobaczymy — nie należy lekceważyć. Droga ta wiedzie przez eksperymentowanie na grupach zróżnicowanych, tak jak zróżnicowane są szersze zbiorowości, przez prowadzenie eksperymentów w różnych warunkach społecznych: przestrzennych i czasowych, a nade wszystko przez pomysłową aranżację, manipulacje i pomiary eksperymentalne tak, by badani traktowali sytuacje stworzone przez eksperymentatora jako sytuacje rzeczywiste i stosownie do tego przekonania się zachowywali. Osiągnięcie trafności zewnętrznej jest więc bardziej rzeczą procedur niż schematów.

6. SCHEMATY EKSPERYMENTALNE

Wykład schematów eksperymentalnych stosowanych lub tylko stosowalnych w badaniach społecznych obejmie dwie grupy schematów. Pierwsza grupa, wyrastająca z tradycji Millowskiej, obejmie schemat klasyczny, oparty wprost na kanonie jedynej różnicy oraz jego rozmaite przekształcenia. Schematy te są schematami jednego bodźca, w dodatku zwykle uję- tego dwuwartościowo. Schematy, które znajdą się w drugiej grupie, wypracowane w latach dwudziestych przez Fishera, zrywają z tymi ograniczeniami. Pozwalają one badać równoczesny wpływ wielu zmiennych, wpływ zmiennych wielowartościowych oraz równoczesny wpływ wielu zmiennych wielowartościowych. Nazwiemy je schematami wielowariantowymi.

6.1. KLASYCZNY SCHEMAT EKSPERYMENTALNY I JEGO PRZEKSZTAŁCENIA

Przegląd schematów eksperymentalnych dla dwóch powodów dobrze będzie zacząć od schematu zwanego klasycznym. Schemat ten w świadomości potocznej zajmuje miejsce szczególne: badania na nim oparte bywają utożsamiane z eksperymentowaniem po prostu, a on sam bywa uważany za schemat idealny; oczywiście, oba te przekonania są z gruntu niesłuszne. Ważniejszy jest wzgląd drugi — dydaktyczny. Oto bowiem istnieje cała wielka rodzina schematów eksperymentalnych, które są różnorodnymi przekształceniami schematu klasycznego i przez odniesienie doń dadzą się łatwo uporządkować, zrozumieć i ocenić[48].

6.1.1. Klasyczny schemat eksperymentalny.

Klasyczny schemat eksperymentalny w swej najprostszej postaci wymaga dwu grup ujednoliconych pod istotnymi względami, a zwłaszcza pod względem zmiennej zależnej. Grupy te mogą składać się z osób, zbiorowości lub instytucji; w szczególnym przypadku liczba jednostek w każdej z grup może równać się jedności. W obu grupach wykonuje się pretest: mierzy się wstępnie przynajmniej zmienną zależną, po czym jedną z nich — eksperymentalną — poddaje się działaniu bodźca, a drugą — kontrolną — zostawia się bez bodźca. Po odpowiednim czasie wykonuje się posttest: grupy mierzy się ponownie. Różnica między zmianą zmiennej zależnej w jednej grupie, a zmianą stwierdzoną w drugiej grupie stanowi miarę wpływu bodźca, gdyż był on jedyną różnicą między grupami.

Jeśli grupę eksperymentalną oznaczymy literą *E*, kontrolną *K*, bodziec — *X* a pretest i posttest literą P z odpowiednimi indeksami, to zapis schematu klasycznego będzie wyglądał następująco:

$$\begin{array}{llll} E & P_1 & X & P_2 \\ K & P_3 & & P_4 \end{array}$$

Wpływ bodźca będzie wyrażać różnica

[48] Wykład tych schematów ma charakter podręcznikowy i mało oryginalny. Do jego przygotowania wykorzystaliśmy rozmaite dostępne nam podręczniki metodologii badań społecznych i monografie poświęcone eksperymentowaniu. Ogólna koncepcja tego wykładu i bardzo wiele stwierdzeń szczegółowych pochodzi z renomowanej monografii D. T. Campbella i J. C. Stanleya Experimental and quasi-experimental designs for research. Praca ta ukazała się po raz pierwszy w r. 1963 i od tego czasu nikt lepszego wykładu schematów eksperymentalnych nie przedstawił. Autorowie rozdziałów poświęconych eksperymentowaniu w podręcznikach metodologii wyrażają natomiast zwykle wyrazy uznania dla pracy Campbella i Stanleya i ograniczają się do zreferowania jej zasadniczych idei. My również nie zamierzamy być oryginalni — wszelka próba oryginalności miałaby tylko jeden skutek pewny a byłoby nim obniżenie wartości wykładu.

$$d = (P_2 - P_1) - (P_4 - P_3)$$

Przykładem eksperymentu opartego na najprostszej wersji schematu klasycznego jest pochodzące z początku lat trzydziestych badanie Dodda nad wpływem upowszechniania wiedzy sanitarnej na poziom higieny na wsi[49]. Grupę eksperymentalną stanowiło kilkadziesiąt rodzin w pewnej wsi syryjskiej. Grupą kontrolną była pewna liczba rodzin w innej wsi o podobnych warunkach geograficznych, demograficznych, historycznych, religijnych, o tym samym poziomie wykształcenia, warunkach rekreacyjnych i sanitarnych. Wsie oddalone były od siebie tak, by upowszechniana wiedza nie dotarła do wsi kontrolnej. Obie wsie zmierzono na skali higieny — wyniki były podobne. Wieś eksperymentalną poddano 2-letniemu szkoleniu w zakresie higieny, prowadzonej przez objazdową ekipę lekarską. Gdy po dwu latach obie wsie zmierzono ponownie, okazało się, że — wbrew oczekiwaniom eksperymetatora — w obu nastąpiła mniej więcej taka sama poprawa poziomu higieny. Wynik taki może wskazywać. na nieskuteczność ocenianego programu. Dodd zinterpretował go inaczej: jego zdaniem brak różnic między wsią eksperymentalną i kontrolną mógł się wziąć z braku zupełnej izolacji wsi kontrolnej od wpływów programu eksperymentalnego.

Schemat klasyczny rzadko stosowany bywa w tak prymitywnej postaci. Zazwyczaj w eksperymentach na nim opartych nie porównuje się po prostu bodźca z brakiem bodźca, ale jedną wartość bodźca z drugą. Tak np. porównuje się wpływ zwiększenia udziału pracowników w podejmowaniu decyzji organizacyjnych na wydajność i zadowolenie z pracy z wpływem zmniejszenia tego udziału[50], skuteczność argumentacji jednostronnej i argumentacji przedstawiającej także kontrargumenty[51], wpływ demokratycznej i autokratycznej atmosfery w grupie na całość życia grupowego[52] itd. W eksperymentach tego rodzaju każda z grup eksperymentalnych stanowi kontrolną grupę dla drugiej. Wprowadza się też, choć nie zawsze grupę kontrolną w ścisłym sensie, taką, na którą nie oddziaływa się żadnymi bodźcami. Jest ona grupą kontrolną dla obydwu, czy też dla wszystkich grup eksperymentalnych. Dla porównania wpływu dwu bodźców nie jest cna niezbędna, konieczna jest natomiast dla bezwzględnej oceny siły każdego z tych bodźców. Eksperymenty tego rodzaju nazywa się niekiedy porównawczymi. Ciekawym przykładem takiego badania jest eksperyment terenowy wykonany w 1935 r. przez Hartmanna w celu porównania siły wpływu propagandy emocjonalnej i racjonalnej na decyzje wyborcze[53]. Hartmann wykorzystał wybory do władz lokalnych w Allentown (USA). Wyodrębnił on w tym mieście dwie grupy obwodów wyborczych, porównywalne ze sobą ze względu na gęstość zaludnienia, status społeczno-ekonomiczny, wartość majątku i nawyki wyborcze mieszkańców; pozostałe obwody potraktował jako grupę kontrolną. Wśród mieszkańców obu grup rozprowadził specjalnie opracowane ulotki wzywające do głosowania na bardzo mało w Allentown popularną partię socjalistyczną, przy czym w jednej grupie ulotki te odwoływały się do emocji, a w drugiej do rozumu. Pretest stanowiły wyniki wyborów poprzednich, a posttest — obecnych. Odpowiednie porównania wykazały, że największy przyrost głosów na socjalistów wystąpił w obwodach poddanych propagandzie emocjonalnej, mniejszy w obwodach, na które skierowano propagandę racjonalną, a najmniejszy w obwodach kontrolnych.

[49] Zob. F. S. Chapin, Experimental designs is sociological research, New York 1955, s. 55 - 7.

[50] N. C. Morse, E. Reimer, The experimental change of a major organizational variable, w: W. M. Evan (ed.), Organizational experiments, New York 1971..

[51] C. I. Hovland, A. A. Lumsdaine, F. D. Sheffield, Experiments on mass communication, Princeton 1949, rozdz. 8.

[52] K. Lewin, R. Lippitt, Eksperymentalna metoda badań nad autokracją i demokracją w: J. N. Kosiński, Socjologia amerykańska, (oprac.) Nowy York [1962].

[53] G.W. Hartmann, A field experiment on the comparative effectiveness of „emotional" and „rational" political leaflets in determining election results, w: R.L. Rosnow, E. J. Robinson (eds.), Experiments in persuasion, New York 1067.

Klasyczny schemat eksperymentalny doskonale pozwala radzić sobie ze wszystkimi omawianymi wyżej czynnikami zakłócającymi wpływ bodźca. Wstępne ujednolicenie grupy eksperymentalnej i kontrolnej wyklucza możliwość wyjaśnienia zaobserwowanych różnic między tymi grupami różnicami początkowymi, bo tych nie było. Ujednolicenie warunków eksperymentu pozwala kontrolować pozostałe czynniki uboczne, gdyż działają one jednakowo na obie grupy. Gdyby zaszły istotne zmiany zewnętrzne, to odbiłyby się w równym stopniu na wielkości zmiany zmiennej zależnej w obu grupach. Podobnie rzecz miałaby się wtedy, gdyby następowały zmiany samorzutne lub istniał pewien trend ogólny. Jeśli pretest wpływałby na zmienną zależną, to również w obu grupach. W obu grupach jednakowe są narzędzia i procedury pomiarowe, to i jednakowe byłyby ich zmiany. Gdyby grupy wybrano ze względu na krańcowe wartości zmiennej zależnej, to w obu nastąpiłaby regresja ku średniej. Wreszcie, ewentualna śmiertelność miałaby jednakowo selektywny charakter w obu grupach. O tym, czy zaobserwowane różnice mogą być wytłumaczone niekontrolowanymi okolicznościami przypadkowymi, łatwo przekonałby odpowiedni test istotności. Dodajmy, że efektywna kontrola czynników kontaminujących w schemacie klasycznym nie wymaga uprzedniego przewidzenia ich pojawienia się — schemat ten równie dobrze kontroluje zakłócenia znane, jak i takie, których eksperymentator nawet się nie domyśla.

Odpowiedni dobór grupy eksperymentalnej i kontrolnej oraz jednakowe ich traktowanie jest więc sposobem eliminacji hipotez konkurencyjnych przez sprowadzenie sytuacji do kanonu jedynej różnicy. Czynniki obecne zarówno tam, gdzie skutek wystąpił, jak i tam, gdzie go nie było, nie mogą być jego przyczynami. Tą może być tylko czynnik obecny tam, gdzie był skutek, *a* nieobecny tam, gdzie go nie było. Warunek ten spełnia tylko bodziec i tylko on może być przyczyną. Czy nią faktycznie jest, to zależy od staranności kontroli tj. od tego, czy faktycznie grupy były jednakowe przed badaniem i czy faktycznie kontakt z bodźcem był jedyną różnicą w warunkach, w jakich się znajdowały w czasie eksperymentu. To zaś nie jest już sprawą schematu, lecz wiedzy, pomysłowości i sumienności eksperymentatorów, którzy ten schemat wypełniają treścią swoich procedur.

Klasyczny schemat eksperymentalny ma również swoje wady i ograniczenia. Trudno niekiedy spełnić jego wymagania — dobrać dwie ujednolicone względem siebie grupy, zapewnić im jednolite. — oprócz bodźca — warunki i móc zmierzyć je w dwóch, rozdzielonych wprowadzeniem bodźca momentach. Schemat klasyczny nie pozwala na badanie szeregu ważnych zagadnień, a jeśli już, to za cenę dużych nakładów pracy. Jako schemat jednego bodźca nie pozwala śledzić wpływu zmiennych przybierających więcej niż dwie wartości. Z tego samego powodu nie można na jego podstawie badać równoczesnego wpływu wielu zmiennych, a jest to zagadnienie kapitalne, zwłaszcza jeśli wziąć pod uwagę interakcje tych zmiennych. Jako schemat dwóch tylko pomiarów nie pozwala badać zmienności następstw bodźców w czasie: ich trwałości, słabnięcia i nasilania się. Jako schemat jednej tylko grupy eksperymentalnej i jednej grupy kontrolnej nie pozwala badać uwarunkowań wpływu bodźca, istotnych dla oceny możliwości uogólnień stwierdzonych zależności.

Świadomość tych trudności sprawia, że schemat klasyczny poddawany jest rozmaitym przekształceniom. Jedne z nich zubożają go o pewne elementy. Inne Wzbogacają go o elementy nowe. Jeszcze inne zubożając go pod pewnymi względami, wzbogacają go pod drugimi. Modyfikacje te mają na celu albo ominięcie trudności stosowania schematu klasycznego albo przystosowanie go do rozwiązywania bardziej skomplikowanych zagadnień niż te, do jakich tak znakomicie się nadaje.

6.1.2. Schemat z dwukrotnym pomiarem w jednej grupie

Niekiedy badacz nie może oprzeć eksperymentu na schemacie klasycznym, gdyż nie dysponuje grupą kontrolną. Może to wynikać z bardzo specyficznego charakteru grupy eksperymentalnej — nie sposób wtedy znaleźć dla niej odpowiednika. Może też brać się to stąd, iż nie sposób ograniczyć zasięgu bodźca do pewnego tylko fragmentu zbiorowości. Zdarza się to zwłaszcza w eksperymentach terenowych, prowadzonych w zbiorowościach i na instytucjach naturalnych.

W takiej sytuacji schemat klasyczny ulega okrojeniu o grupę kontrolną. Powstaje schemat z dwukrotnym pomiarem w jednej grupie.

$$E \qquad P_1 \qquad X \qquad P_2$$

Miarą wpływu bodźca jest różnica *d = P2 - P1*

Przykładem eksperymentu opartego na tym schemacie jest badanie Malewskiego[54]. Sprawdzał on hipotezę, że ludzie o pozytywnym obrazie własnej osoby po podjęciu decyzji zwiększają atrakcyjność wybranej możliwości, a ludzie o obrazie negatywnym — zmniejszają ją. Badani układali pewne przedmioty według ich atrakcyjności, po czym mogli wybrać dla siebie jeden z nich i wreszcie układali wszystkie przedmioty ponownie. Stwierdzone w ten sposób zmiany w atrakcyjności przedmiotów wybranych i odrzuconych korelowano z obrazem własnej osoby. Eksperyment ten nie przewidywał grupy kontrolnej.

Dyskutowany schemat ma złą reputację wśród metodologów a trochę tylko lepszą wśród badaczy. Nie ma w tym nic dziwnego, jeśli przyjrzeć się założeniom, jakich on wymaga.

Schemat „pretest — posttest" czyni stan grupy eksperymentalnej w chwili wykonywania pretestu sytuacją kontrolną dla posttestu. Zakłada, że gdyby nie interwencja eksperymentatora, to wartość zmiennej zależnej nie zmieniłaby się. Innymi słowy, że bodziec jest jedynym czynnikiem, który może wyjaśnić stwierdzoną różnicę w wynikach pomiarów.

Ujmując rzecz bardziej szczegółowo rozważany schemat zakłada, że: 1. W środowisku zewnętrznym nie działają równolegle z bodźcem inne czynniki wpływające na zmienną zależną. 2. Wewnątrz badanych układów nie dzieje się nic takiego, co wpływa na zmienną zależną; w szczególności nie istnieje ogólny trend, którego fragmentem mogłaby być stwierdzona zmiana. 3. Pretest nie wpływa na wynik post- testu. 4. W przypadku grup krańcowych nie wystąpiła regresja wartości zmiennej zależnej ku średniej. 5. Narzędzia pomiaru są niewrażliwe na czas. 6. W składzie grupy nie wystąpiły selektywne zmiany.

Są to bardzo mocne założenia. Ciężar dowodu spoczywa na eksperymentatorze, a jest on tym dotkliwszy, że sformułowane są one w sposób negatywny, a dowieść negację jest rzeczą trudną. 'Tylko przyjąwszy te założenia można konkluzywnie stwierdzić, że zaobserwowana różnica jest dziełem bodźca. Aby ten wniosek obalić, wystarczy dowieść, że którekolwiek z wyszczególnionych założeń nie zostało w eksperymencie spełnione. Wtedy zmiana zostanie wyjaśniona konkurencyjnie: nie przez bodziec, a przez czynnik go kontaminujący.

W tym schemacie eksperymentalnym bardziej niż w innych widać, że z istotności uzyskanych różnic nie wynikają żadne wnioski pozytywne dotyczące ich pochodzenia.

Czy wynika stąd, że eksperymenty oparte na dyskutowanym schemacie są zupełnie bezwartościowe jako metoda dowodu przyczynowego? Bynajmniej! Schemat ten jest ryzykownym w stosunku odwrotnie proporcjonalnym do stopnia uzasadnienia jego założeń. Jeżeli w konkretnej sytuacji badawczej eksperymentator zasadnie przyjmie, że wskazane wyżej zakłócenia nie przekraczają rozmiarów tolerowanych albo gdy potrafi on odizolować od nich grupę eksperymentalną, to będzie miał prawo sądzić z różnicy wyników pomiarów o wpływie bodźca. Z sytuacją taką mamy do czynienia szczególnie w krótkotrwałych eksperymentach laboratoryjnych — w badaniach tych tworzenie grup kontrolnych, choć możliwe — jest niejednokrotnie zbędne.

6.1.3. Schematy szeregów czasowych

Słaby schemat z dwukrotnym pomiarem w jednej grupie, uzyskany w wyniku zubożenia schematu klasycznego, można wzmocnić elementami, których w sche-

[54] A. Malewski. Pozytywny i negatywny obraz własnej osoby a proces podejmowania decyzji, w: O nowy kształt nauk społecznych, Warszawa 1974, s. 237.

macie klasycznym nie było, a mianowicie dodatkowymi pretestami i posttestami. W wyniku tego powstaje schemat pojedynczego szeregu czasowego.

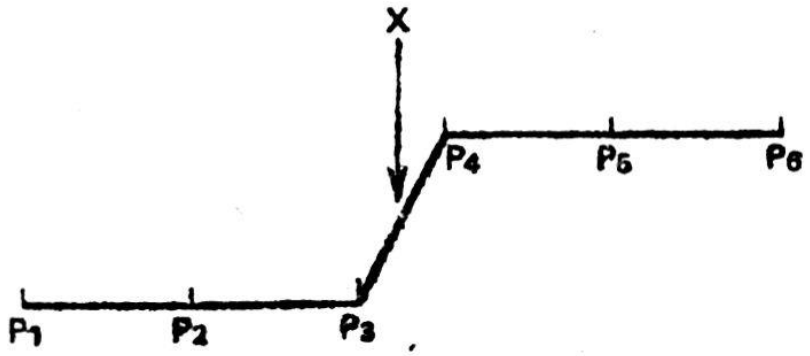
$$E \qquad P_1 \quad P_2 \quad P_3 \quad X \quad P_4 \quad P_5 \quad P_6$$

Wpływ bodźca widoczny jest w nieciągłości wyników pomiarów dokonywanych przed wprowadzeniem bodźca i po jego wprowadzeniu.

Na schemacie pojedynczego szeregu czasowego oparte były pionierskie badania prowadzone na przełomie pierwszego i drugiego dziesięciolecia w przemyśle brytyjskim przez British Industrial Fatigue Research Board[55]. W jednym z tych eksperymentów w sierpniu 1919 r. w pewnej fabryce butelek zmieniono dziesięciogodzinny dzień pracy na dziewięciogodzinny i badano skutki tej zmiany w godzinowej wydajności pracy. Badacze dysponowali wskaźnikami wydajności pracy dla kilkunastu miesięcy poprzedzających zmianę i kilkunastu po niej następujących.

Schemat pojedynczego szeregu czasowego pozwala kontrolować wiele zakłóceń, z którymi schemat poprzedni nie daje sobie rady. Dwa hipotetyczne szeregi wyników przedstawione niżej umożliwią ilustrację tej tezy.

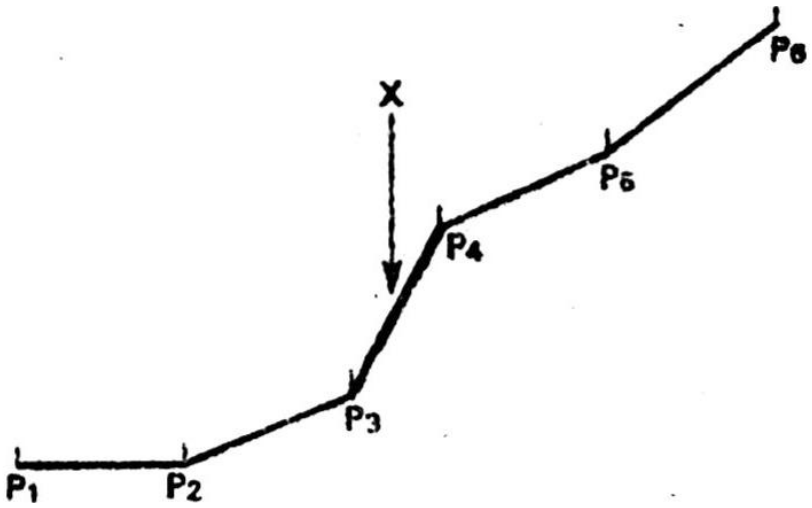
W pierwszym wypadku rozciągnięcie pomiarów eksperymentalnych wstecz i w przód pozwala wyeliminować zmiany wewnętrzne (trend), pretest, zmiany



Rysunek 1

narzędzi pomiarowych oraz selektywne ubytki jako źródła konkurencyjnych w stosunku do bodźca wyjaśnień różnicy P_4-P_3. Nie ma bowiem powodu, dla którego miałyby one działać tylko wtedy, gdy wprowadzono bodziec, skoro nie działały ani przedtem

[55] Zob. D. T. Campbell, From description to experimentation: interpreting trends as quasi-experiments, w: Ch. W. Harris (ed.), Problems in measuring change, Madison 1963.



Rysunek 2

ani potem. W drugim przypadku szereg czasowy chroni przed groźbą bezpodstawnego przypisania całej różnicy P_4-P_3 wpływowi bodźca. Pokazuje bowiem, że pewne zmiany występowały systematycznie również przed i po interwencji. Różnica ta może więc być częściowo spowodowana działaniem wymienionych wyżej czynników zakłócających.

Pojedynczy eksperyment oparty na schemacie przerwanego szeregu czasowego nie pozwala jednak kontrolować zmian zewnętrznych. Aby to czynić, trzeba - byłoby przeprowadzić ten sam eksperyment w różnych, podejrzewanych o zakłócający wpływ, warunkach. Jeśli powtarzając eksperyment uzyska się ten sam wynik, można będzie wykluczyć także wpływ zmian zewnętrznych, przynajmniej tych, które były zróżnicowane.

Lepszą metodą jest dobranie do eksperymentalnego szeregu czasowego — szeregu kontrolnego. W wyniku tego powstaje schemat wielokrotnego szeregu czasowego:

E	P_1	P_2	P_3	**X**	P_4	P_5	P_6
K	P_7	P_8	P_9		P_{10}	P_{11}	P_{12}

Miarą wpływu bodźca jest różnica między zmianą w szeregu eksperymentalnym a zmianą w szeregu kontrolnym.

Jest to bardzo silny schemat i dobrze się on nadaje do eksperymentów w instytucjach. W badaniach tych trudno dobrać do instytucji eksperymentalnych identyczne instytucje kontrolne; instytucje są złożone i nie mają swych bliźniaczych odpowiedników Ale ponieważ możliwe są w nich pomiary wielokrotne, zwykle systematycznie prowadzone przez same te instytucje, to można niedostatki doboru skompensować analizą szeregów czasowych.

Schematy szeregów czasowych posiadają istotne ograniczenie. Można je stosować tylko do badania takich środowisk, w których rozmaitego rodzaju pomiary, takie jak testy czy oceny, są rzeczą naturalną oraz do badania takich układów, w których niezależnie od eksperymentu prowadzi się systematyczne pomiary ważnych zmiennych; szkoła i fabryka są przykładami takich instytucji. Ograniczenie to bierze się z niebezpieczeństwa możliwych zmian zachowania pod wpływem pomiarów i pod wpływem świadomości, że jest się przedmiotem badania. Zmiany tego rodzaju ograniczyłyby możliwość uogólniania uzyskanych wyników na sytuacje nieeksperymentalne.

6.1.4. Schemat X grupą kontrolną bez pretestu

Wpływ bodźca w schemacie klasycznym wyrażała różnica między zmianą w grupie eksperymentalnej a zmianą w grupie kontrolnej. Jeśli jednak obie grupy były starannie dobrane i pretest dał jednakowe wyniki w obu, to wpływ bodźca równie dobrze wyraża różnica między wynikami posttestów w obu grupach. Zatem, znajomość konkretnej wartości zmiennej zależnej przed manipulacją wcale nie jest konieczna do zmierzenia wpływu bodźca — wystarczy wiedzieć, że wartość ta była jednakowa w obu grupach. Przekonanie to można uzyskać przez losowy dobór jednostek do poszczególnych grup albo przez celowe wyrównanie grup ze względu na wartości tych zmiennych, o których zakładamy, że są skorelowane ze zmienną zależną.

W schemacie klasycznym można więc zmniejszyć liczbę pomiarów i uzyskać w efekcie schemat z grupą kontrolną bez pretestu.

E	X	P_1
K		P_2

Różnica $d = P1 - P2$ jest miarą wpływu bodźca.

Schemat ten jest bardzo popularny w badaniach społecznych. O popularności tej decydują jego zalety:

1. Pretest nie zawsze jest potrzebny, czasem przeciwnie — czyniłby on badanie bardziej praco- i czasochłonnym, nie miałby sensu lub wręcz uniemożliwiałby cały eksperyment. Schemat z grupą kontrolną bez pretestu bierze te względy pod uwagę.

Asch badając wpływ nacisku grupowego na wydawane sądy, porównywał decyzje osób poddanych naciskowi z decyzjami osób podejmujących decyzje samodzielnie, grupy te dobrał uprzednio tak, by były one podobne pod ważnymi względami osobowościowymi[56]. Potocka-Hoserowa badając wpływ lęku na sprawność umysłową dobrała dwie grupy uczniów równie inteligentnych, po czym kazała jednej z nich wykonywać pewne zadanie w warunkach stresu lękowego, a drugiej w warunkach wolnych od stresu[57]. W obu przypadkach eksperymentatorzy nie mierzyli zmiennej zależnej przed wprowadzeniem bodźca: w pierwszym z eksperymentów pretest, pomagając badanym domyślić się celu badania, odebrałby mu sens, a w drugim byłby całkiem zbędny, wszak związek zmierzonej inteligencji ze sprawnością wykonywania zadań umysłowych ma charakter definicyjny. Pretest trzeba wykonywać raczej wtedy, gdy zmienną niezależną jest jakaś cecha, którą wszyscy w jakimś stopniu posiadają. Tam zaś, gdzie bodziec ma wywołać pewne konkretne zachowanie mało prawdopodobne bez niego, pretest nie ma sensu — można co najwyżej zmierzyć zmienne istotne dla tego rodzaju zachowania.

2. Schemat z grupą kontrolną bez pretestu świetnie pozwala radzić sobie ze wszystkimi czynnikami zakłócającymi wpływ bodźca. Zmiany zewnętrzne i wewnętrzne (rozwój samorzutny, trend) odbijają się w równym stopniu na obu grupach. Z wpływem pretestu samego i zmianą narzędzi pomiarowych w ogóle nie mamy do czynienia, gdyż pomiar jest jednorazowy. Obie grupy są ujednolicone ze względu na wartość zmiennej zależnej i (lub) zmiennych istotnych. Ujednolicenie składu grup i warunków, w jakich funkcjonują zapewnia, że ewentualne nielosowe ubytki z grup będą dokonywać się według tych samych zasad. Ze względu na siłę kontroli zakłóceń bodźca schemat dyskutowany w niczym nie ustępuje schematowi klasycznemu, a niekiedy go przewyższa.

3. Swoją przewagę nad schematem klasycznym wykazuje dyskutowany schemat najlepiej wtedy, gdy badacz — stwierdziwszy zależność — przechodzi do jej

[56] S. Asch, Nacisk grupy na modyfikacje i wypaczenia sądów, w: A. Malewski (red.) Zagadnienia psychologii społecznej, Warszawa 1962.

[57] A. Potocka-Hoserowa, Uczenie społeczne, lęk przed niepowodzeniem i sprawność intelektualna, „Studia Socjologiczne" 1963, nr 2.

uogólniania na sytuacje pozaeksperymentalne. Otóż schemat klasyczny (jak zresztą i pozostałe dotąd omówione schematy) posiada istotną wadę ograniczającą" uogólnianie wyników badań na nim opartych. Umożliwia on taki wpływ pretestu na badanych, który warunkuje działanie bodźca na zmienną zależną. Zjawisko to zwykło się nazywać interakcją bodźca z pretestem. Za określeniem tym kryje się kilka rodzajów wpływów pretestu na wpływ bodźca: (a) uwrażliwianie na bodziec, (b) inhibicja wpływu bodźca i (c) informowanie badanych o eksperymentalnym charakterze sytuacji, w jakiej się znajdują; może są i inne kanały wpływu.

(a) Pretest uwrażliwia badanych na bodziec eksperymentalny: wzmaga zainteresowanie nim, wpływa na jego percepcję, ułatwia zauważenie tych aspektów bodźca, które są istotne dla badania. Wpływ ten jest szczególnie istotny w badaniach zmiany postaw za pomocą złożonych wielowarstwowych przekazów symbolicznych takich jak filmy czy audycje radiowe.

Pretest kieruje uwagę badanych na te strony przekazu, których wpływ chce się badać. W efekcie, na badanych protestowanych bodziec działa inaczej — silniej niż na badanych niepretestowanych i jeśli nawet na początku badania pierwsi z nich byli reprezentatywni dla pewnej zbiorowości, to pod wpływem protestu reprezentatywnymi być przestają.

(b) Pretest może też działać jako inhibitor bodźca — czynnik osłabiający jego wpływ. Zjawisko to występuje w badaniach zmiany postaw przy pomocy argumentacji perswazyjnej. Otóż pretest wymaga od badanych wyrażenia konkretnej postawy a przez to przyczynia się do tego, że utwierdzają się oni w swych przekonaniach. Stąd badani protestowani w mniejszym stopniu niż niepretestowani ulegną perswazji. Istnienie tego zjawiska sugerowali już Howland i jego współpracownicy w interpretacji głośnych eksperymentów z zakresu komunikacji masowej[58]. Zbadał go zaś w serii eksperymentów Lana[59]. Wykazał on m. im., że inhibicyjny wpływ protestu na wielkość zmiany postaw występuje wyraźnie wówczas, gdy argumentacja zawiera argumenty za i przeciw, a nie zaznacza się wtedy, gdy argumentacja jest jednostronna. W pierwszym wypadku wpływ protestu będzie miał okazję wystąpić u każdego z badanych, gdyż każdy z nich napotyka w przekazie argumenty wymierzone przeciw jego przekonaniom, a w drugim — wpływ ten wystąpić może tylko u części badanych.

(c) Pretest jest jednym *ze* źródeł, z którego badani w eksperymentach prowadzonych w środowiskach naturalnych mogą się domyślać eksperymentalnego charakteru bodźców, którym zostają poddani, a w eksperymentach laboratoryjnych mogą się domyślać faktycznego celu badania lub przynajmniej sformułować sobie własne jego interpretacje. Rzecz jasna, pretest nie jest jedynym źródłem tej wiedzy, zwłaszcza w eksperymentach laboratoryjnych. Aranżacja badania, zachowanie się eksperymentatora, sam charakter bodźców pełnić mogą tę samą, co pretest rolę; obszerniej powiemy o tym dalej. Wiele badań empirycznych wykazało, że wiedza ta w istotny sposób wpływa na reakcje badanych na bodźce eksperymentalne.. Badani świadomi celów eksperymentatora lub choćby tylko eksperymentalnego charakteru sytuacji mogą stawiać bodźcowi większy opór' niż osoby, które takiej wiedzy nie mają. Wiele częściej jednak dostosowują oni swe zachowanie do oczekiwań imputowanych eksperymentatorowi. Zjawiska te, tak samo jak i uwrażliwiający wpływ pretestu, utrudniają uogólnianie stwierdzonych zależności na jednostki, które nie posiadają takiej wiedzy i takich postaw, jak badani.

O istotności pretestu dla wpływu bodźca łatwo się przekonać. Wystarczy porównać wyniki uzyskane w eksperymencie opartym na schemacie klasycznym z wynikami tego samego eksperymentu, opartego na schemacie z grupą kontrolną bez pretestu. Sugestia ta leży u podstaw schematu Solomona.

Schemat z grupą kontrolną bez pretestu wyklucza zarówno uwrażliwiający i inhibicyjny wpływ pretestu na bodziec, jak też informowanie badanych przez pretest

[58] Hovland i in., op. cit., s. 308 - 13.

[59] R. E. Lana, Inhibitory effects of a pretest on opinion change, w: R. L. Rosnow, E. J. Robinson (eds.), Experiments in persuasion, New York 1967.

o fakcie i celach eksperymentu. Czyni to w sposób najskuteczniejszy z możliwych — przez rezygnację z pretestu. Wyniki eksperymentów na nim opartych zyskują tym samym na ogólności.

Schemat z grupą kontrolną bez pretestu posiada też i wady, których schemat klasyczny nie ma.

1. Założenie o wyrównaniu w grupie eksperymentalnej i kontrolnej wartości zmiennej zależnej przed wyprowadzeniem bodźca ma charakter hipotetyczny. Nie osłabia to siły dowodu przyczynowego, gdy grupy te dobiera się losowo, zwłaszcza gdy są to grupy liczne. W innych jednakże wypadkach trafność tego założenia zależy od wartości teorii wiążącej wielkości zmierzonych zmiennych istotnych z niemierzonymi wielkościami zmiennych zależnych; teoria ta z ujednolicenia pierwszych pozwala wnosić o ujednoliceniu drugich. Z uwagi na częstą słabość takich teorii, dobór losowy winien być stosowany wszędzie, gdzie jest on tylko możliwy.

2. Brak pretestu uniemożliwia obserwację zmian na poziomie jednostek, z których składają się grupy — umożliwia tylko obserwację zmian na poziomie grup. Ma to ujemne następstwa dla opisu i wyjaśnienia zmian wywołanych przez bodziec.

a) Nie można zbadać procesu, który doprowadził do obserwowanych różnic w rozkładach zmiennej zależnej na poziomie grup. Nie można śledzić zmian wewnątrz grup, a tylko produkt końcowy tych zmian.

b) Nie docenia się rozmiarów zmiany. Jeśli jednostkowe zmiany wewnątrz grup mają kierunki sprzeczne, to znoszą się i na poziomie grupy widoczna jest tylko ich wypadkowa.

c) Poważnie ograniczone są możliwości badania uwarunkowań zróżnicowania wpływu bodźca. Dlaczego jedne jednostki zmieniły się pod wpływem bodźca, a drugie nie? Szukanie odpowiedzi na to ważne pytanie jest standardowym krokiem w analizie wyników eksperymentu, następującym zaraz po ustaleniu ogólnych rozmiarów i jednostkowego zróżnicowania zmiany. Na gruncie schematu z grupą kontrolną bez pretestu pytanie to często musi pozostać bez odpowiedzi, bowiem nie wiadomo, które jednostki zmieniły się, a które nie. Jest to najpoważniejsza wada dyskutowanego schematu, dotkliwa zwłaszcza wtedy, gdy chcemy przewidzieć reakcje jednostek na bodźce w sytuacjach naturalnych; ograniczona znajomość modyfikatorów wpływu bodźca osłabia siłę przewidywania.

Czynnikiem, który może różnicować, wpływ bodźca jest początkowa wartość zmiennej zależnej. Tymczasem takie klasyczne zagadnienie, jak zależność wpływu perswazji od początkowej rozbieżności między postawami nadawcy a odbiorcy, w ogóle nie da się badać na gruncie tego schematu.

Pozostałe czynniki trzeba podzielić na dwie grupy: na takie, które jak wiek czy wykształcenie nie są podatne na wpływ bodźca i na takie, które jak różne postawy są nań podatne, tak samo jak zmienna zależna. Modyfikującą siłę pierwszych można badać przez podział badanych wedle ich wartości i ustalenie różnic w wartościach zmiennej zależnej między utworzonymi w ten sposób grupami. Na badanie modyfikującej siły drugich na gruncie tego schematu nie ma jednak sposobu.

W schemacie z grupą kontrolną bez pretestu, podobnie jak i w schemacie klasycznym rzadko używa się grupy eksperymentalnej z bodźcem i kontrolnej bez bodźca. Zwykle są to dwie grupy eksperymentalne, stanowiące nawzajem dla siebie kontrolę, każda z innym wariantem bodźca.

Schemat z grupą kontrolną bez pretestu, jak zresztą wszystkie schematy wymagające grup kontrolnych, zakładał taki przydział jednostek do grup eksperymentalnych i kontrolnych, który uniezależnia kontakt z- bodźcem lub jego pewnym wariantem od uprzednich własności jednostek, zwłaszcza zaś wartości zmiennej zależnej i zmiennych istotnych. Uniezależnienie to najlepiej zapewniał przydział losowy. Po przydzieleniu jednostek do poszczególnych grup, eksperymentator poddawał działaniu bodźca wszystkie jednostki w grupie eksperymentalnej. To więc, czy grupa jest czy nie jest eksperymentalną nie zależy od tego, czy poddano ją bodźcowi. Jest dokładnie odwrotnie — grupę najpierw wybiera się jako eksperymentalną, a potem się nią manipuluje.

Inaczej jest w eksperymentach o schemacie łudząco podobnym do schematu z grupą kontrolną bez pretestu, a mianowicie w takich, które polegają na wprowadzeniu bodźca do pewnej zbiorowości — zwykle naturalnej — i porównywaniu tych, którzy się z nim zetknęli z tymi, którzy się z nim nie zetknęli. Pierwszych uznaje się za grupę eksperymentalną, drugich — za kontrolną, a z różnic między nimi wnosi się o wpływie bodźca. Bardzo to naiwne rozumowanie.

Wprowadzenie do zbiorowości bodźca eksperymentalnego, takiego jak przekaz masowy nie jest równoznaczne z poddaniem jego działaniu każdego członka tej zbiorowości; dociera on do pewnej tylko jej części, niekiedy bardzo małej. Gdy sam schemat eksperymentu nie określa kategorii jednostek, które będą poddane i nie poddane manipulacji, pojawia się pytanie: Jakie czynniki warunkują kontakt z bodźcem? Czy są to czynniki losowe czy też autoselekcja dokonuje się ze względu na zmiene istotne dla wpływu bodźca?

Namysł teoretyczny sugeruje a badania empiryczne potwierdzają, że o wystawianiu się na działanie bodźców decydują uprzednio posiadane cechy jednostek," przedmiotowo związane z bodźcami. Teoria dysonansu poznawczego każe oczekiwać, że w reakcji na środki masowego przekazu ludzie będą szukać informacji zbliżonych do swych przekonań, a unikać i odrzucać informacje z nimi sprzeczne. W eksperymentalnych badaniach wpływu kampanii informacyjnej Star i Hughes ustalili, że „ludzie, którzy mieli uprzednio przychylne postawy, są tymi, którzy zwracają uwagę na kampanię"[60] a inni badacze wykazali, że ci, co świeżo kupili pewien towar, częściej zwracają potem uwagę na jego reklamę;[61] badań takich wykonano zresztą wiele. Co więcej, między kontakt z bodźcem a udzielenie badaczowi informacji o kontakcie interweniuje pamięć, równie jak postrzeganie selektywna.

Z ustaleń tych wynika, że stwierdzone korelacje między kontaktem z bodźcem a posiadaniem cech, które ten bodziec miał kształtować, wcale nie oznaczają wpływu bodźca na te cechy, lecz mogą oznaczać wpływ dokładnie odwrotny.

Mechanizm autoselekcji, polegający na wystawianiu się na bodziec osób o postawach bliskich propagowanym, ma ujemne następstwa także dla wielkości wpływu bodźca na tych, co się nań wystawili. Efekt sufitowy sprawi, że nawet bodziec, który wywołałby zmianę w grupie reprezentującej całą skalę wartości zmiennej zależnej, w tej szczególnej grupie wpływu nie okaże — na zmianę nie ma już w niej miejsca.

Stąd w eksperymentach na grupach dobranych losowo obserwuje się silniejszy wpływ bodźców perswazyjnych niż w eksperymentach, w których grupa eksperymentalna wyłania się na zasadzie autoselekcji, a reszta zbiorowości stanowi grupę kontrolną[62]

Eksperymenty dyskutowanego rodzaju są bardziej podobne do korelacyjnych badań sondażowych niż do eksperymentów z grupą kontrolną bez pretestu. Cechą, która je różni od pierwszych a upodabnia do drugich jest celowa, motywowana względami poznawczymi manipulacja. Różni je zasadniczo metoda kontroli. W eksperymencie z grupą kontrolną bez pretestu kontrolę osiąga się przez ujednolicenie grup przed manipulacją. W eksperymencie dyskutowanym, tak jak i w sondażu, osiąga się ją przez dokonywaną *er post* analizę wielozmiennową. Z trzech warunków dowodu przyczynowego: współwystępowania zdarzeń, rzeczywistego a nie pozornego związku między nimi oraz ich porządku czasowego schemat dyskutowany pozwala stwierdzić tylko dwa pierwsze, przy czym drugi z wiele większym ryzykiem niż schemat z grupą kontrolną bez pretestu; problem następstwa czasowego zmiennych pozostaje na jego gruncie nierozwiązany.

[60] S. A. Star, H. M. Hughes, Report on an educational campaign: the Cincinnati plan for the United Nations, „American Journal of Sociology" 1950, vol. 55, nr 4.

[61] Zob. A. Malewski, O zastosowaniach teorii zachowania, w: O nowy kształt nauk społecznych, Warszawa 1974, s. 404.

[62] Zob. C. I. Hovland, Reconciling conflicting results derived from experimental and survey studies of attitude change, w: E. P. Hollander, R. G. Hunt (eds.), Classic contributions to social psychology, New York 1972.

6.1.5. Schematy z dwukrotnym pomiarem w różnych grupach (bez i z grupami kontrolnymi)

Omawiając schemat z dwukrotnym pomiarem w jednej grupie, powiedzieliśmy, że stosuje się go m. in. wtedy gdy nie sposób do grupy eksperymentalnej dobrać grupy kontrolnej, z powodu jej specyfiki lub niemożliwości ograniczenia zasięgu bodźca. Ale bywa, że i ten schemat zastosować trudno. Badani mogą się nie godzić na ponowny kontakt w celu wykonania posttestu lub grupa eksperymentalna może cechować się dużą płynnością składu. Powiedzieliśmy też, że schemat ten jest bardzo wrażliwy na zakłócenia a interakcja pretestu z bodźcem ogranicza ogólność wyników eksperymentów na nim opartych. Ze wskazanymi trudnościami praktycznymi, niektórymi zakłóceniami i interakcją protestu z bodźcem można sobie dać radę, gdy zastosuje się pewną modyfikację tego schematu, a mianowicie schemat z dwukrotnym pomiarem w różnych grupach. Schemat ten pozoruje schemat poprzedni, gdyż każdy z pomiarów wykonany jest na innej grupie.

Oto zapis tego schematu. Nawias oznacza, że X-a w grupie *E1* może nie być w ogóle, a jeśli już pojawi się on (wskutek braku ograniczeń jego zasięgu), to i tak w analizie nie czyni się z tego żadnego użytku.

$$E_1 \quad P_1 \quad (X)$$
$$E_2 \quad (X) \quad P_2$$

Miarą wpływu bodźca jest różnica d = *P2* - P*1*

O grupach trzeba założyć, że są one ujednolicone ze względu na początkową wartość zmiennej zależnej i zmiennych istotnych, że gdyby *E2* zmierzono po raz pierwszy razem z *E1,* to wynik pomiaru byłby tylko nieistotnie różny od P*1*. Losowy dobór jednostek do grup usprawiedliwia takie założenie.

Przykładem badania opartego na tym schemacie jest eksperymentalna kampania informacyjna przeprowadzona w latach 1947/48 w Cincinnati (USA), zmierzająca do zwiększenia zainteresowania mieszkańców tego miasta sprawami międzynarodowymi[63]. W kampanii tej użyto różnych publicznych środków przekazu, nie można więc było podzielić miasta na dzielnice eksperymentalne i kontrolne, tak jak zrobił Hartmann w opisanym wyżej badaniu propagandy racjonalnej i emocjonalnej, gdy posłużył się ulotkami przynoszonymi do każdego domu oddzielnie. W celu określenia rozmiaru zmian w zainteresowaniach wykonano dwa sondaże, przed i po kampanii, każdy na innej reprezentacyjnej próbie mieszkańców miasta. Okazało się, że w okresie między sondażami wzrosło zainteresowanie tylko stosunkami ze Związkiem Radzieckim i groźbą nowej wojny; zainteresowanie innymi problemami nie zmieniło się. Zasadniczym problemem interpretacji było wyjaśnienie tej zmiany i ocena wpływu kampanii. Star i Hughes referując wynik eksperymentu zauważają, że w czasie eksperymentu nasilała się „zimna wojna" i raczej ten fakt a nie kampania informacyjna spowodował zaobserwowaną zmianę; brak wzrostu zainteresowania innymi sprawami popiera ich wniosek.

Schemat z dwukrotnym pomiarem w różnych grupach przewyższa schemat z dwukrotnym pomiarem w jednej grupie zarówno pod względem trafności wewnętrznej jak i zewnętrznej. Będąc równie jak tamten wrażliwy na zakłócający wpływ bodźca, historię i trend, ma nad nim tę przewagę, że z listy czynników zakłócających eliminuje wpływ pierwszego pomiaru na drugi (w referowanym eksperymencie miałby on postać efektu panelowego) oraz ubytki. Istotna dla ogólności ustalonych związków interakcja pretestu z bodźcem nie jest na gruncie tego schematu możliwa. Zastosowany w tym schemacie pośredni pomiar grupy eksperymentalnej przed wprowadzeniem bodźca spotkamy i w innych schematach.

[63] Zob. Star, Hughes, op. cit.

Schemat rozważany ustępuje za to schematowi z dwukrotnym pomiarem w jednej grupie w zakresie możliwości analizy wyników. Ponieważ za każdym razem mierzy się inną grupę, to traci się możliwość analizy na poziomie jednostkowym, a więc badania przepływów i uwarunkowań zmiany. Schemat ten wystarcza tam, gdzie badacza interesują tylko zmiany na poziomie globalnym.

Podstawowe niebezpieczeństwo w interpretacji wyników eksperymentu opartego na tym schemacie jest takie samo, jak w schemacie z dwukrotnym pomiarem w jednej grupie: jest nim możliwość wzięcia zmian spowodowanych przez inne wydarzenia lub będących fragmentem trendu za skutki bodźca. Eksperyment w Cincinnati pokazuje to niebezpieczeństwo w sposób dobitny. Podobnie jak w schemacie z jedną grupą trzeba też sobie z nim radzić. Aby skontrolować wpływ zmian zewnętrznych, trzeba powtórzyć badanie w zróżnicowanych pod tym względem sytuacjach i gdy uzyska się te same wyniki, będzie to dowód, że zmiany te są nieistotne dla stwierdzonej zależności. Aby skontrolować zmiany wewnętrzne, zwłaszcza trendy, trzeba dokonać kilku kolejnych protestów, za każdym razem na innej grupie ujednoliconej względem grupy eksperymentalnej. Porównanie wyników tych pretestów pozwoli stwierdzić, czy w ogóle istnieje jakiś trend, którego fragmentem mogłaby być różnica między posttestem a ostatnim z pretestów. Schemat ten byłby takim odpowiednikiem schematu przerwanego szeregu czasowego, jakim dla schematu z dwukrotnym pomiarem w jednej grupie jest schemat z dwukrotnym pomiarem w różnych grupach.

Dyskutowany schemat można rozwinąć i wzbogacić przez dodanie doń grup kontrolnych. Tak jak pretestu i posttestu w grupie eksperymentalnej dokonywało się na oddzielnych, lecz ujednoliconych grupach, tak tu na oddzielnych grupach dokonuje się pomiarów -kontrolnych. Założeniem tego schematu jest ujednolicenie wszystkich grup biorących udział w eksperymencie.

E_1	P_1	(X)	
E_2		(X)	P_2
K_1	P_3		
K_2			P_4

Miarą wpływu bodźca jest różnica

$$d = (P_2 — P_1) — (P_4 — P_3).$$

Schemat ten w miarę wzrostu pewności ujednolicenia grup zbliża się w sile wnioskowania do schematu klasycznego, a przewyższa go tym, że unika interakcji pretestu z bodźcem. Jeśli chodzi o możliwości analizy danych, to nie pozwala on na badanie zmiany na poziomie jednostkowym wraz ze wszystkimi wynikającymi stąd ujemnymi konsekwencjami, o których była już wyżej mowa. Schemat ten zaproponowany został przez Campbella i Stanleya Nie znamy eksperymentu na nim opartego.

6.1.6. Schemat Solomona

Poważną wadą schematu klasycznego było dopuszczanie przezeń interakcji pretestu z bodźcem, zjawiska częstego zwłaszcza w eksperymentach nad zmianą postaw i uczeniem się, ograniczającego ogólność ustalonych zależności. Wada ta była tym dotkliwsza, że eksperyment klasyczny kontrolował wszystkie kontaminacje bodźca oraz umożliwiał bardziej skomplikowane niż inne schematy analizy wyników. Słabość schematu klasycznego skutecznie przezwyciężał schemat z grupą kontrolną bez pretestu: ponieważ nie wymagał on protestu, wykluczał tym

samym jego interakcję z bodźcem. Jednak cena płacona przez niego za eliminację interakcji była wysoka: poważne zubożenie możliwości analizy wyników. Inny, mniej kosztowny, choć bardziej pracochłonny sposób kontroli interakcji polega nie na uniemożliwieniu jej, lecz na jej zmierzeniu. Drogą tą poszedł Solomon[64].

Solomon uzupełnił klasyczny, -schemat eksperymentalny o dodatkowe dwie grupy. Nie przeprowadza się w nich pretestów, lecz tylko posttesty, przy czym jedną z nich poddaje się temu samemu bodźcowi, co i grupę eksperymentalną. W ten sposób obok grupy eksperymentalnej istnieją trzy grupy kontrolne. Mógłby ktoś powiedzieć, że schemat Solomona składa się z dwu grup eksperymentalnych i dwu kontrolnych, lecz nie byłoby to ścisłe, gdyż druga z grup poddanych bodźcowi pełni przede wszystkim funkcje kontrolne wobec pierwszej.

Zapis tego schematu wygląda następująco:

E	P_1	X	P_2
K_1	P_3		P_4
K_2		X	P_5
K_3			P_6

Jeśli wszystkie grupy dobrano losowo, to w grupach dodatkowych *K2 i K3* można obejść się bez pretestów; gdyby go przeprowadzono dałby on wyniki takie same jak pretesty wykonane w dwu pierwszych grupach *E* i *K1*. Zamiast więc mierzyć, wylicza się wynik hipotetycznego prestestu jako średnią dwóch pretestów wykonanych *Ph = ½(P1 + P2)*.

Zmiana *d1 = P2 - P1* w grupie eksperymentalnej *E* jest wynikiem działania bodźca, pretestu, ujętych łącznie zmian zewnętrznych i samorzutnych oraz interakcji pretestu z bodźcem. Zmiana *d2 = P4 - P3* w grupie kontrolnej *K1* wynika z wpływu pretestu oraz zmian wewnętrznych i samorzutnych. Zmiana *d3 = = P5 — Ph* w grupie kontrolnej K2 wyraża wpływ bodźca oraz zmian zewnętrznych i samorzutnych, a zmiana *d4 = P6 - Ph* w grupie kontrolnej *K3* — tylko wpływ owych zmian.

Schemat Solomona pozwala na łatwe wyliczenie nie tylko czystego wpływu bodźca, ale i innych, zanieczyszczających jego wpływ czynników. Jak już powiedzieliśmy, wpływ zmian zewnętrznych i samorzutnych wyraża różnica *d4* w grupie *K3*. Wpływ czystego bodźca to różnica między zmianą *d3* w *K2* a zmianą *d4* w *K3*. Wpływ samego pretestu wyraża różnica między zmianą *d2* w *K1* a zmianą *d4* w *K3*.

Co ważniejsze, schemat Solomona pozwala mierzyć interakcję pretestu z bodźcem. Jeśli zmiana zaobserwowana w grupie eksperymentalnej wywołana przez wszystkie czynniki działające łącznie jest różna od sumy zmian wywołanych przez poszczególne czynniki oddzielnie, to znaczy, że interakcje zachodzące między tymi czynnikami warunkują wpływ bodźca. Może tu wchodzić w grę zarówno wpływ zmian zewnętrznych i samorzutnych na działanie bodźca, jak też i wpływ pretestu. Pierwszy rodzaj interakcji jest zwykle mało prawdopodobny, wobec czego różnica

$$I = d_1 - [(d_3 - d_4) + (d_2 - d_4) + d_4] = d_1 - (d_2 + d_3 - d_4)$$

jest miarą interakcji bodźca z protestem.

Schemat Solomona ma też postać skróconą. Tam, gdzie nie ma zanieczyszczających wpływy bodźca zmian zewnętrznych i samorzutnych, różnica *d4* w grupie *K3* równałaby się zeru. Zatem można się obyć bez tej grupy. Interpretacja różnic w poszczególnych grupach, wyliczenie czystego wpływu bodźca i pretestu oraz

[64] R. L. Solomon, An extension of control group design, „Psychological Bulletin" 1919, vol. 46, nr 2.

wyliczenie i interpretacja ich interakcji staje się wówczas odpowiednio prostsza. Różnica *d1* w grupie eksperymentalnej wyraża wpływ bodźca, pretestu oraz ich interakcji, różnica *d2* — wpływ samego pretestu a różnica *d3* — wpływ czystego bodźca. Zatem różnica *I' = d1 — (d2 + d3)* jest miarą interakcji.

Przykładem eksperymentu opartego na schemacie Solomona jest jedno z serii badań wykonanych przez Lanę mających na celu ustalenie, w jaki sposób pretest w eksperymentach nad zmianą postaw przez perswazję warunkuje wpływ bodźca[65]. Bodźcem w tym eksperymencie był przekaz jednostronny, zawierający tylko argumenty za wiwisekcją, a nie zawierający kontrargumentów. Lana podzielił losowo stu kilkudziesięciu studentów na 4 grupy, po czym zmierzył je odpowiednim kwestionariuszem i poddał perswazji, tak jak wymaga tego schemat Solomona. Jedynym czynnikiem zmieniającym postawy okazał się bodziec; interakcja pretestu z bodźcem nie wystąpiła. Ale w innych eksperymentach, o czym już zresztą wspominaliśmy, Lana dowiódł, że gdy argumentacja zawiera argumenty za i przeciw, pretest istotnie zmniejsza wpływ bodźca.

Rzut oka na pełny schemat Solomona przekonuje, że stanowi on połączenie schematu klasycznego ze schematem z grupą kontrolną bez pretestu. Te dwa schematy różnią się istnieniem pretestu. Ponieważ wpływ pretestu bezpośrednio na zmienną zależną lub posttest jest kontrolowany na gruncie schematu klasycznego, to jedyną różnicą między tymi schematami jest możliwość interakcji pretestu z bodźcem. Zatem, różnica między wynikami dwu eksperymentów, z których każdy oparto na innym schemacie, będzie miarą interakcji.

$$I = (d_1 - d_2) - (d_3 - d_4)$$

Równanie to jest równoważne poprzedniemu.

Porównania takie przeprowadzała grupa Hovlanda w badaniach skuteczności propagandy[66]. Badacze ci stwierdzili, że propagandowy efekt filmu zmierzony na podstawie schematu klasycznego był mniejszy od efektu ustalonego na podstawie schematu bez pretestu. Różnicę tę wyjaśnili oni pewną postacią interakcji pretestu z bodźcem, specyficzną dla schematu klasycznego, a mianowicie wpływem tendencji do udzielania odpowiedzi spójnych *(consistency reaction)* na pomiar wpływu bodźca. Tendencja ta może wynikać stąd, że badani zapamiętawszy swoje odpowiedzi w preteście, świadomie uzgadniają z nimi owe odpowiedzi w poście. Może też brać się stąd, że pretest zmuszając badanych o niesprecyzowanych opiniach do udzielania sprecyzowanych odpowiedzi, krystalizuje opinie i uogólnia odpowiedzi. Tendencja ta nie tylko zaciemnia wpływ bodźca, z czym łatwo można byłoby sobie poradzić przy pomocy grupy kontrolnej, lecz także zmniejsza go, a do wykrycia tego wpływu schemat klasyczny nie wystarcza.

Zadaniem schematu Solomona nie jest badanie wpływu bodźca oczyszczonego z kontaminacji i wolnego od uwarunkowań — do tego wystarczyłaby… z powodzeniem schemat z grupą kontrolną bez pretestu, nieporównanie odeń prostszy. Schemat Solomona służy przede wszystkim do pomiaru interakcji z bodźcem i dalej, do oceny wartości schematu klasycznego w badaniu określonej dziedziny. Jeśli okaże się, że interakcji nie ma, to znaczy to, iż schemat klasyczny wolno stosować dalej bez obaw o niekontrolowane na jego gruncie zakłócenia, a możliwości analizy i interpretacji wyników eksperymentu klasycznego są — jak pisaliśmy wyżej — znacznie większe niż w eksperymencie z grupą kontrolną bez pretestu. Schemat Solomona jest więc przede wszystkim narzędziem badań metodologicznych.

6.1.7. Schemat bez pretestu i bez grupy kontrolnej.

Klasyczny schemat eksperymentalny, a także schemat z grupą kontrolną bez pretestu umożliwiał stwierdzenie wszystkich trzech warunków koniecznych do

[65] Lana, op. cit.
[66] Hovland i in., op. cit., dodatek C.

dowodu przyczynowego: współwystępowanie zdarzeń, ich następstwa czasowego oraz rzeczywistego charakteru związku między nimi. Schemat z dwukrotnym pomiarem bez grupy kontrolnej pozwalał stwierdzić tylko dwa pierwsze warunki. Istnieje jednak jeszcze uboższy schemat — taki, który pozwala ustalić tylko współwystępowanie zdarzeń. Mowa o schemacie bez pretestu i bez grupy kontrolnej.

$$E \qquad X \qquad P$$

Miarą wpływu bodźca jest wynik posttestu.

Jeżeli ze współwystępowania dwu zdarzeń — wprowadzonego bodźca i obserwowanego zjawiska polegającego na przyjęciu przez zmienną zależną pewnej wartości, chce się wnosić, że pierwsze z nich jest przyczyną drugiego, to trzeba założyć, że obserwowane zjawisko pojawiło się dopiero po wprowadzeniu bodźca i że nie mogło ono zostać spowodowane przez inne zdarzenie niż bodziec. Schemat wymagający tak mocnych założeń jest zbyt naiwny, by się nad nim dłużej zatrzymywać — posiada on wszelkie wady niekontrolowanego badania korelacyjnego.

Niemniej i te założenia niekiedy, choć rzadko się spełnia i na prymitywnym schemacie wykonuje się znakomite eksperymenty. Dziać się może tak wtedy, gdy efektem manipulacji ma być pewne wyraźne i specyficzne zachowanie, nieprawdopodobne bez manipulacji, a nie np. wzrost stopnia posiadanej cechy. Przykładem niech będzie pierwsze z głośnej serii badań Milgrama nad posłuszeństwem zgodnym z rozkazem, a sprzecznym z sumieniem[67].

Milgram polecił badanym — przekonanym, że biorą udział w eksperymentach nad wpływem karania na uczenie się, uczyć innych badanych pewnych słów, a za błędy karać ich wstrząsami elektrycznymi o rosnącym natężeniu. W rzeczy samej, przedmiotem zainteresowania byli sami „nauczyciele", w przewodach generatora wstrząsów nie płynął prąd, „uczniowie" odgrywali tylko role ofiar, a celem badania było ustalenie, jak daleko w zadawaniu bólu protestującej ofierze posuną się badani na polecenie eksperymentatora. Przed rozpoczęciem doświadczenia eksperci przewidywali, że prawie wszyscy badani odmówią posłuszeństwa eksperymentatorowi, a usłuchają głosu sumienia. Wynik stanowił ogólne zaskoczenie: ⅔ badanych posunęło się do zadawania bliźnim maksymalnych, groźnych dla życia (450 V) wstrząsów elektrycznych, tylko dlatego, że otrzymało takie polecenie, nie poparte żadnymi sankcjami, a sprzeczne z etyką elementarną; badani wykazywali przy tym oznaki skrajnego napięcia psychicznego, rzadko spotykanego w laboratorium psychospołecznym.

W referowanym eksperymencie Milgram nie stosował ani grupy kontrolnej, ani pomiaru wstępnego. Nie sposób uzasadnić potrzeby tworzenia grupy kontrolnej: dlaczegóżby badani, którym nie polecono zadawać innym bolesnych wstrząsów elektrycznych, zadawać je mieli? Nie sposób też wyobrazić sobie, jak mógłby wyglądać i czego powinien dotyczyć pretest: zadawania cierpienia bez rozkazu?

Dyskutowany schemat, choć niezmiernie zawodny jako podstawa wnioskowania przyczynowego, pełni pożyteczną rolę dydaktyczną. „Eksperyment z jednokrotną obserwacją i bez grupy kontrolnej pozwala pokazać, jaką rolę przy wnioskowaniu z eksperymentu odgrywa wielokrotna obserwacja i jaką rolę odgrywa badanie więcej niż jednej zbiorowości"[68]

6.1.8. Schematy do badania wpływów zmieniających się w czasie

Eksperymentatorzy w badaniach społecznych, a zwłaszcza psychospołecznych mają skłonność do ustalenia tylko bezpośrednich następstw zastosowanych bodźców albo też do stosowania tylko takich bodźców, których następstwa dają się obserwować po niedługim czasie.

---

[67] S. Milgram. Behavioral study of obedience, „Journal of Abnormal and Social Psychology'' 1963, vol. 67, nr 4. Pełny raport: Obedience to authority; an experimental view, New York 1974.

[68] J. Karpiński, Weryfikacja tez o związkach przyczynowych w badaniach eksperymentalnych „Studia Socjologiczne" 1977, nr 3.

Bez wątpienia oba te rodzaje badań są cenne, a często również i wystarczające do sformułowania zespołu praw rządzących pewną dziedziną. Co więcej, niektóre schematy eksperymentalne oparte są na założeniu przejściowego charakteru wywołanych zmian. I to jeszcze zanotować trzeba, że niejednokrotnie taki właśnie charakter następstw manipulacji eksperymentalnych jest warunkiem, dopuszczalności eksperymentowania.

Niemniej jednak wskazane ograniczenie powoduje istotne następstwa negatywne. Odstęp czasowy między wieloma zjawiskami połączonymi więzią przyczynową jest długi. Wiele zmiennych — następstw manipulacji definuje się przez odniesienie do czasu; stabilność względnie niestabilność składu grup czy postaw niech będą przykładami. Wreszcie, dynamika rozmaitych procesów wywołanych eksperymentalnie wyraża się właśnie zmianami częstotliwości i natężenia pewnych zjawisk w czasie. W każdym z tych kontekstów dokonywanie tylko jednorazowego pomiaru, w dodatku — zaraz po wprowadzeniu i zadziałaniu bodźca doprowadziłoby do zubożenia i zniekształcenia obrazu badanych zjawisk. Niebezpieczeństwa te nie są tylko hipotetycznymi konstrukcjami metodologów.

Carl Hovland, wytrawny badacz zmian postaw zauważył, że typowe eksperymenty w tej dziedzinie prowadzą do przeceniania wpływu bodźców perswazyjnych. W eksperymentach tych w zasadzie efekt mierzy się właśnie zaraz po wprowadzeniu bodźca. W badaniach sondażowych, w których efekty przekazów mierzy się zwykle po dłuższym upływie czasu, uzyskuje się wnioski świadczące o mniejszej sile perswazyjnej przekazów[69]. Oba rodzaje wniosków są wartościowe, rzecz w tym, by prowadzić takie badania, które pozwolą na chwytanie skutków tak bezpośrednich jak i odległych.

Naturalnym sposobem badania zmian skutków manipulacji w miarę upływu czasu byłoby badanie oparte na opisanych już schematach, uzupełnione o dodatkowe posttesty. Metoda ta jest jednak stosowalna tylko tam, gdzie pomiary mogą mieć charakter dyskretny i same nie wpływają na zmianę. W pozostałych wypadkach, schematy uzupełnia się o nowe grupy eksperymentalne i kontrolne, w których posttesty dokonuje się później niż w grupach podstawowych.

W badaniach nad zmianą postaw schematy takie stosowała grupa Hovlanda[70]. Badacze ci użyli dla równoczesnego zmierzenia bezpośrednich i odległych efektów propagandowych filmów, schematu będącego połączeniem dwóch schematów klasycznych, z których jeden służył uchwyceniu skutków bezpośrednich, a drugi — odległych.

E_1	P_1	X	P_2
K_1	P_3		P_4
E_2	P_5	X	P_6
K_2	P_7		P_8

Schemat ten można poddawać rozmaitym modyfikacjom, podobnie jak i schemat klasyczny; jedną z takich modyfikacji przedstawimy poniżej. Wadą tego schematu jest brak możliwości kontroli zakłóceń występujących w okresie między pierwszym a drugim posttestem. Może się zdarzyć, że wpływ określonego wydarzenia w tym okresie mylnie weźmiemy za interesujący nas wpływ czasu *per se*. Drugą wadą jest dopuszczenie interakcji pretestu z bodźcem; jej wpływ na zmienną zależną może również zmieniać się w czasie.

Schematy tego rodzaju pozwalają odkrywać nie tylko słabnięcie — w miarę upływu czasu — skutków manipulacji eksperymentalnych. Bywa i przeciwnie czas

[69] Zob. Hovland, op. cit.
[70] Zob. Hovland i in., op. cit., rozdz. 7.

wzmacnia niektóre zmiany. Zjawisko to ma nawet swą nazwę w teorii i badaniach zmiany postaw — efekt śpiącego (sleeper *effect).* „Efekt śpiącego — definiują Capon i Hulbert — można zaobserwować wtedy, gdy jednostka zgadza się bardziej z przekazem perswazyjnym po upływie dłuższego czasu od prezentacji bodźca niż bepośrednio po niej a przy tym końcowy pomiar zgody, porównany z pretestem, wykazuje zmianę w kierunku propagowanym”[71]. Przykładem efektu śpiącego jest wzrost — w miarę upływu czasu — skuteczności perswazji wywieranej przez źródło oceniane przez odbiorcę negatywnie. Ponieważ źródło zapomina się szybciej niż treść, a im wyższa ocena źródła, tym wpływ większy, to zmiana narasta w miarę zapominania źródła[72]

Capon i Hulbert proponują dobry schemat dla eksperymentów, w których chciałoby się badać przejściowe, opóźnione wpływy bodźców.

$$
\begin{array}{llll}
E_1 & X & P_1 & \\
K_1 & & P_2 & \\
E_2 & X & & P_3 \\
K_2 & & & P_4 \\
K_3 & P_5 & &
\end{array}
$$

Ważną innowacją tego schematu w stosunku do schematu stosowanego przez grupę Hovlanda jest — jak widać — rezygnacja z pretestów we wszystkich czterech grupach i zastąpienie ich pretestem pośrednim wykonywanym na dodatkowej grupie kontrolnej; wszystkie grupy muszą być więc dobrane losowo. Unika się przez to kontaminującego wpływu pretestu oraz — co ważniejsze — interakcji pretestu z bodźcem.

6.1.9. Schemat z powtarzaniem bodźca w jednej grupie

Bardzo wysoki stopień kontroli eksperymentalnej możemy uzyskać nawet wtedy, gdy nie dysponujemy grupą kontrolną, analogiczną pod istotnymi względami do grupy eksperymentalnej i jesteśmy zdani tylko na porównania w jednej grupie. Ta sama grupa może pełnić równocześnie funkcje grupy eksperymentalnej i kontrolnej, a przedmiotem porównania mogą być nie grupy, ale stany grupy w różnych momentach. Jeżeli bodziec jest tego rodzaju, że jego następstwa mają charakter przejściowy i szybko znikają lub wręcz trwają dotąd tylko dopóki działa bodziec, to możemy wprowadzać do tej samej grupy bodziec wielokrotnie, co pewien czas i wielokrotnie mierzyć zmienną zależną w warunkach działania bodźca i przy jego braku. Możemy też wprowadzać na przemian różne warianty bodźca i mierzyć zmienną zależną w warunkach działania każdego z nich. Momenty wprowadzania bodźca względnie wariantów bodźca winny być wybrane losowo dla tych samych powodów, dla których losowo przydziela się jednostki do grup eksperymentalnych i kontrolnych.

Oto zapis tego schematu dla najprostszej sytuacji bodźcowej (X, X):

$$XP_1 \quad \bar{X}P_2 \quad XP_3 \quad \bar{X}P_4 \quad XP_5 \quad \bar{X}P_6 \ \ldots \ XP_{n-1} \quad \bar{X}P_n$$

Miarą wpływu bodźca jest różnica średnich dwu serii pomiarów, z których jedną wykonano wtedy, gdy bodziec działał, a drugą wtedy, gdy nie działał: P_1, P_3, $P_5 \ldots P_{n_1}$ i P_2, P_4, P_6, P_n.

[71] N. Capon, J. Hulbert, A sleeper effect — an awakening, „Public Opinion Quarterly” 1973, vol. 37, nr 3.

[72] Zob. Th. M. Newcomb i in., Psychologia społeczna, Warszawa 1970, s. 121.

W jednym z eksperymentów opartych na tym schemacie Kerr badał wpływ muzyki na wydajność pracy w fabryce[73]. Porównywał np. średnią wydajność pracy w próbie 56 dni, w których nadawano muzykę z próbą 51 dni, w których muzyki nie nadawano; próby te pobrał losowo z kilku miesięcy. W innym eksperymencie porównywał on wpływ różnych rodzajów muzyki na wydajność pracy na trzech próbach 14- dniowych. Zakładać on musiał, że wpływ muzyki lub pewnego rodzaju muzyki na wydajność ustaje do dnia, w którym mierzy się wydajność pracy bez akompaniamentu lub przy innym akompaniamencie. Inaczej, wnioski nie byłyby trafne.

Schemat dyskutowany posiada silę bliską sile schematu klasycznego. Nie pozwala on wytłumaczyć stwierdzonej różnicy żadnymi ze znanych czynników ubocznych, gdyż wszystkie one są kontrolowane. Ponieważ bodziec wprowadza się i pomiarów dokonuje się wielokrotnie i w momentach wybranych losowo, jest mało prawdopodobne, by jakieś zakłócające zmiany w środowisku pojawiały się tylko wtedy, gdy wprowadzono bodziec i znikały wraz z nim. Natomiast takie zakłócenia, jak trend, wpływ pretestu i zmiany narzędzi pomiarowych dotyczą jednakowo obu serii pomiarów. Regresja, selektywny dobór i selektywne ubytki uniemożliwione są przez losowy dobór momentów wprowadzania bodźca.

Schemat z powtarzaniem bodźca w jednej grupie dopuszcza jednak nowy, niespotykany w dotychczas przedstawionych schematach, czynnik zakłócający i warunkujący wpływ określonego bodźca. Jest nim inny bodziec. Nie zakończony do chwili wprowadzania bodźca *X2* wpływ bodźca *X1,* stwarza niebezpieczeństwo dwojakiego rodzaju. Po pierwsze, nakładanie się wpływów obu bodźców utrudnia określenie wpływów każdego z nich oddzielnie. Po wtóre, jeden bodziec warunkuje wpływ drugiego. Drugie ze zjawisk, będące interakcją „bodziec × bodziec", nazywa się w metodologii efektem przeniesienia *(carry over effect).* Egzemplifikacjami tych zjawisk w eksperymentach Kerra mogłyby być: długotrwały wpływ muzyki jednego rodzaju na wydajność pracy, powodujący nakładanie się wpływów różnych rodzajów muzyki na siebie oraz uzależnienie wpływu muzyki każdego z rodzajów od urozmaicenia warunków pracy przez ciągłe zmiany akompaniamentu. Trudno nam powiedzieć, czy te dające się pomyśleć niebezpieczeństwa faktycznie zagrażały trafności wewnętrznej i zewnętrznej wyników uzyskanych przez Kerra.

6.1.10. Schemat rotacyjny

Schemat rotacyjny stosuje się wówczas, gdy wprawdzie dysponuje się grupami porównawczymi, ale nie można ich wyrównać na tyle, by różnice początkowe między nimi nie mogły tłumaczyć różnic w wynikach posttestów. Z taką sytuacją ma się często do czynienia w eksperymentach na zbiorowościach naturalnych, niewielkich a przy tym zróżnicowanych, takich jak np. klasy szkolne. Stąd schemat rotacyjny bywa stosowany w eksperymentach pedagogicznych.

Schemat rotacyjny polega na naprzemiennym wprowadzaniu bodźca do każdej z dwu grup, gdy tymczasem druga grupa służy do kontroli; można też porównywać efekty dwu wariantów bodźca.

$$\begin{array}{lll} E_1 & XP_1 & XP_2 \\ E_2 & XP_3 & XP_4 \end{array}$$

Miarą wpływu bodźca jest różnica

[73] W. A. Kerr, Experiments on the effect of music on factory production, „Applied Psychology Monographs" 1945. nr 5; podajemy za Campbellem i Stanleyem.

$$d = \frac{1}{2}\,(P_1 + P_4) - \frac{1}{2}\,(P_2 + P_3).$$

Przykładem eksperymentu opartego na schemacie rotacyjnym jest badanie wykonane przez Nozeta w celu porównania wartości dydaktycznej dwu wersji filmu oświatowego: niemej i dźwiękowej[74], Badacz ten wybrał dwie klasy szkolne, tylko mniej więcej równe. W pierwszej fazie eksperymentu w jednej z klas wyświetlał film niemy a w drugiej z klas film dźwiękowy; w drugiej fazie postępował odwrotnie. Po sprawdzeniu wiadomości związanych z treściami filmu i dokonaniu odpowiednich porównań okazało się, że film niemy 'był bardziej skuteczny niż dźwiękowy — zmuszał on uczniów do większej aktywności.

Schemat rotacyjny jest schematem pomysłowym i prostym. Jego główną zaletą jest uniemożliwienie faworyzowania lub upośledzania bodźca przez istniejące różnice między grupami. Jego słabość tkwi w możliwości dotrwania następstw jednego wariantu bodźca do chwili wprowadzania drugiego wariantu. O wynikających stąd konsekwencjach dla trafności wewnętrznej i zewnętrznej eksperymentu pisaliśmy już przy przedstawieniu poprzedniego schematu tj. schematu z powtarzaniem bodźca w jednej grupie.

6.1.11. Schemat cykliczny jako przykład schematów kombinowanych

W dotychczasowym przeglądzie schematów eksperymentalnych często pojawiały się schematy złożone, będące wynikiem połączenia pewnych schematów prostszych. Takim kombinowanym schematem był schemat Solomona a także — służące badaniu bezpośrednich i odległych następstw bodźca — schematy stosowane przez Hovlanda oraz proponowane przez Capona i Hulberta. Schematy te są jednak stosunkowo proste w porównaniu ze schematem, jaki wymyślili Campbell i Stanley. Na fragmencie tego schematu Campbell i McCormack[75] oparli badania nad wpływem szkolenia kadetów na ich postawy wobec różnych aspektów służby wojskowej. Program szkolenia kadetów nie był programem eksperymentalnym, nie odbiera to jednak tej ilustracji siły perwazyjnej.

Programowi szkolenia poddawano co roku nową grupę kadetów, przychodzących do szkoły. Nie można kontrolować jego wpływu ani przez porównanie kadetów ze studentami szkół cywilnych, ani przez dokonanie podziału rocznika na dwie części — eksperymentalną i kontrolną. Szansę kontroli dawał za to fakt, że programowi poddawano co roku nową grupę i że można było przeprowadzać dowolną liczbę pomiarów. Campbell i McCormack opracowali więc następujący plan badania:

I r.	X	P_1			
II r.		P_2	X	P_3	

P1 i P2 wykonano równocześnie; *P1* na grupie kadetów, którzy właśnie ukończyli roczne szkolenie, a *P2* na grupie, która miała je zacząć, *P3* przeprowadzono na drugiej grupie po roku.

Miarę wpływu programu stanowią różnice *P2 - P1* oraz *P3 - P2*. Cały ten schemat powstał w wyniku połączenia schematu z grupą kontrolną bez pretestu ze schematem z dwukrotnym pomiarem bez grupy kontrolnej.

Różnica *P2 — P1* może być jednak wyjaśniona konkurencyjnie, bez odwoływania się do wpływu szkolenia. Można ją wyjaśnić selektywnym doborem: I rocznik

[74] Zob. W. Zaczyński, Praca badawcza nauczyciela, Warszawa 1976, s. 101 - 7.

[75] D. T. Campbell, Th. H. McCormack, Military experience and attitudes toward authority, „American Journal of Sociology" vol. 62, nr 7.

nie musiał być przed rozpoczęciem szkolenia taki sam jak rocznik II w tym samym czasie. Można ją też wyjaśnić dojrzewaniem (w tym miejscu termin ten ma sens nie tylko przenośny): w chwili dokonywania pomiarów *P1* i *P2* rocznik I jest już o rok starszy, bardziej dojrzały od rocznika II. Różnicę *P2 — P1* można wreszcie wytłumaczyć selektywną „śmiertelnością": z rocznika I już niektórzy odpadli a zmienna decydująca o odpadzie mogła być powiązana z mierzonymi postawami. Jednakże różnicy tej nie można wyjaśnić ani niespecyficznymi dla szkoły zmianami zewnętrznymi (dotyczyły one jednakowo obu roczników), ani wpływem pretestu i zmianą narzędzi pomiarowych, bo pomiar był jednorazowy.

Różnica *P3 - P2* sama w sobie nie jest — jak wiadomo — dobrą miarą wpływu bodźca, ale gdyby okazało się, że jest ona równa różnicy *P2 - P1*, to selektywny dobór oraz selektywne ubytki byłyby wykluczone z listy konkurencyjnych wyjaśnień zmiany *P2 - P1*. Niekontrolowany pozostałby tylko jeden czynnik — dojrzewanie.

Campbell i Stanley proponują udoskonalić zastosowany w badaniach kadetów schemat przez losowy podział II rocznika na grupę, którą zbada się przed rozpoczęciem szkolenia i grupę, której się nie zbada oraz przez uwzględnienie kolejnego, III rocznika i zmierzenie go po przyjściu do szkoły. W efekcie uzyskuje się schemat:

I	r.	X	P_1			
II_1	r.		P_2	X	P_3	
II_2	r.			X	P_4	
III	r.				P_5	X

Przez losowy podział II rocznika zyskujemy trzecią miarę wpływu bodźca: *P4 — P2*. Porównanie to pozwala kontrolować zakłócający i interakcyjny wpływ pretestu, w czym przewyższa porównanie *P3 - P2* i zarazem selektywny dobór, w czym przewyższa *P2 - P1*. Dodanie rocznika III i wstępne go zmierzenie dostarcza dodatkowych porównań, kontrolujących wyniki dotychczasowe, np. *P5 - P4* porównuje się *z P2 - P1 a* porównaniem *P5 - P2* zmierzy się różnice w składzie kolejno przychodzących do szkoły kadetów (selektywność doboru).

W dalszym ciągu niekontrolowane pozostaje dojrzewanie, z którym radzić sobie trzeba albo przez porównanie badanych z innymi grupami, jeśli istnieją takie porównywalne grupy, albo przez podział rocznika według wieku, jeśli rocznik jest w tym względzie zróżnicowany. Brak różnic w postawach będzie świadczył w obu wypadkach o nieistotności dojrzewania dla zmiany postaw kadetów Wobec służby wojskowej.

Omówiony schemat jest dobrym narzędziem do badania wpływu instytucji, w których pewien aspekt czy program przedstawiany jest systematycznie coraz to nowej grupie. Szkoła jest taką instytucją. Schemat ten, choć mało elegancki, wykorzystuje do maksimum możliwości eksperymentowania i analizy. Stanowi on dobrą ilustrację pewnej, proponowanej przez Campbella i Stanleya, strategii planowania eksperymentu w sytuacjach nietypowych: trzeba mianowicie tworzyć wiele możliwości oceny bodźca i różnicować je tak, żeby słabe strony jednych porównań uzupełniane były mocnymi stronami drugich. Termin „schematy łatane" użyty jako nazwa budowanych w ten sposób schematów dobrze zdaje sprawę z ich estetycznych i praktycznych właściwości.

6.2. TECHNIKI DOBORU GRUP W EKSPERYMENCIE

W eksperymentach opartych na schematach wymagających więcej niż jednej grupy, grupy powinny być identyczne. Żądanie to, rozumiane dosłownie, jest niewykonalne, gdyż Leibniz uczy, że dwie rzeczy, które mają wszystkie własności wspólne, to jedna i ta sama rzecz. Wobec tego grupy powinny być podobne do siebie pod jak największą ilością względów istotnych. Tak sformułowany postulat jest już wykonalny. Im ściślejsza jego realizacja, tym wnioskowanie przyczynowe pewniejsze.

Oto krótki przegląd technik doboru grup w eksperymentach opartych na dotąd omówionych schematach. Techniki te służą do wyrównywania grup a także do alokacji bodźców. Są one ułożone wedle malejącej roli doboru losowego a rosnącej roli doboru celowego w tych procedurach[76].

Pierwszą z tych technik jest randomizacja. Wymóg ścisłości kazałby mówić tu o kompletnej randomizacji, gdyż sam dobór losowy jest elementem także innych technik doboru grup w eksperymencie, jednakże przyjęte jest mówić po prostu o randomizacji.

Randomizacja polega na losowym podziale pewnej zbiorowości jednostek na poszczególne grupy określone przez schemat eksperymentu. Niekiedy jest ona dwustopniowa: polega na losowym pobraniu wymaganej liczby prób jednostek z pewnej zbiorowości i następnie — na losowym przydzieleniu tych grup poszczególnym wariantom manipulacji; jednym z nich może być brak oddziaływania. Grupy tworzone przez randomizację nie muszą być równoliczne aczkolwiek bez powodu nie powinny być różnoliczne.

Wyniki eksperymentu opartego na randomizacji przedstawia się w postaci tylu zbiorów obserwacji, ile jest grup. Analiza wyników polega na badaniu wielkości różnic między parametrami tych grup, takimi jak proporcje, średnie, wariancje, albo między zmianami parametrów w grupach. Przyjmowane w analizie hipotezy zerowe mówią, że różnice te są równe zeru. Istnieje wiele testów istotności tych różnic, dostosowanych do rodzaju zmiennej losowej, parametrów, mocy użytych skal pomiarowych, wielkości prób oraz tego, czy próby te są zależne czy niezależne.

Randomizacja jako metoda wyrównywania grup i alokacji bodźców odznacza się następującymi zaletami:

1. Umożliwia ona wysoki stopień wyrównania grup względem siebie. Grupy te są wyrównane nie tylko ze względu na zmienną zależną, lecz także ze względu na dużą (teoretycznie rzecz biorąc — nieskończoną) liczbę innych zmiennych: istotnych i nieistotnych, mierzalnych i niemierzalnych, znanych i nieznanych. Wyrównane są rozkłady tych zmiennych oraz ich kombinacje.

2. Randomizacja uniezależnia w wysokim stopniu poddanie jednostki konkretnemu wariantowi bodźca od jej uprzednich cech, a w szczególności od wartości zmiennej zależnej. Losowe przydzielenie jednostek poszczególnym wariantom sprawia też, iż żaden wariant nie będzie ani systematycznie faworyzowany, ani systematycznie upośledzony względem pozostałych przez korelację ze zmiennymi istotnymi wzmacniającymi lub hamującymi jego wpływ.

3. Randomizacja nie wymaga wiedzy o jednostkach składających się na badaną zbiorowość. Przydział jednostek do poszczególnych grup nie zależy od uprzedniego posiadania o nich informacji. Znika w ten sposób jedno ze źródeł obciążenia wyników eksperymentu błędem systematycznym.

4. Randomizacja zwalnia od konieczności wykonywania pretestu, jeśli jego jedynym celem miałoby być zbadanie, czy zmienna zależna ma tę samą wartość we wszystkich grupach. Grupy dobrane losowo z założenia mają ujednolicone rozkłady zmiennych. W ten sposób wyeliminowany zostaje kontaminujący i warunkujący wpływ pretestu na bodziec.

[76] Zagadnienia, o których będziemy mówić dalej, wyłożone są szczegółowo w wielu podręcznikach eksperymentowania, np. w podręcznikach cyt. w przyp. 33.

5. W sytuacjach, w których konieczna jest znajomość wartości zmiennej zależnej przed wprowadzeniem bodźca, randomizacja umożliwia dyskretne protesty pośrednie, dokonywane na dodatkowej grupie kontrolnej. Z sytuacją tą mieliśmy do czynienia np. w schemacie z dwukrotnym pomiarem w różnych grupach, i w jednym ze schematów służących do badania zmieniających się w czasie wpływów bodźca.

6. Zmienne w grupach dobranych losowo mają rozkłady w przybliżeniu normalne, co czyni prostą analizę wyników. Pospolite testy istotności różnic między parametrami prób budowane są dla sytuacji, w których zmienne w tych próbach mają rozkład normalny (np. test t Studenta, test F).

7. Gdy grupy w eksperymencie zostały przed przydzieleniem ich poszczególnym wariantom bodźca wybrane w sposób losowy z szerszej zbiorowości, to wyniki eksperymentu są reprezentatywne dla tej zbiorowości; randomizacja jest więc także metodą podwyższania trafności zewnętrznej eksperymentu a nie tylko trafności wewnętrznej.

8. Randomizacja sprawia, że jednostki składające się na grupy eksperymentalne, reprezentują całą skalę wartości zmiennej zależnej i zmiennych istotnych, co (a) wyklucza regresję wartości zmiennej zależnej ku średniej, (b) pozwala badać uwarunkowania wpływu bodźca przez wartości zmiennych istotnych, jeśli tylko same one nie są podatne na wpływ i (c) podnosi ogólność stwierdzonych zależności (np. efekt sufitowy , nie jest tu możliwy).

9. *Last, but not least,* sama randomizacja jest czynnością prostą.

Zalety randomizacji mogą być wykorzystywane tym lepiej, im większe mają być tworzone przy jej pomocy grupy i im mniej zróżnicowana jest zbiorowość, z której się je losuje. Jeśli zbiorowość owa jest silnie zróżnicowana a grupy mają być niewielkie, randomizacja traci na dokładności.

W wielu dziedzinach badań wiadomo z góry, że pewne cechy jednostek wpływają na zmienną zależną. Nasuwa się więc możliwość zwiększenia dokładności wyrównania grup, jeśli dokładność uzyskana przez randomizację nie byłaby zadowalająca. Trzeba najpierw pogrupować jednostki w zespoły o jednakowych wartościach owych cech istotnych i dopiero wtedy każdy taki zespół podzielić losowo na grupy określone przez schemat eksperymentu. Na tym polega technika zwana po angielsku *matching,* co zwykle tłumaczy się jako „dobór parami", choć para jest tylko najprostszym zespołem tworzonym na podstawie uprzednią wiedzy a dzielonym przez losowanie. Omawiając tę technikę, będziemy — dla prostoty wykładu — mieć na myśli dobór parami; potem przedstawimy jego rozwinięcie.

W wyniku doboru parami i losowego podziału w parach powstają dwie równoliczne grupy, w których każdemu elementowi jednej pary odpowiada analogiczny pod wyrównanymi względami element drugiej grupy, „bliźniaczy odpowiednik" — jak mówią Goode i Hatt. Wynik eksperymentu można przedstawić — tak samo, jak w poprzednim wypadku — w postaci dwu zbiorów obserwacji, oddzielnie dla każdej grupy. Bardziej uzasadnione jest jednak przedstawienie go w postaci jednego zbioru różnic w parach. Analiza polegać będzie wówczas na badaniu średniej tych różnic. Jej wielkość będzie miarą wpływu bodźca. Jeśli zmienna różnic w parach ma rozkład w przybliżeniu normalny. do oceny istotności średniej różnic stosuje się test *t* Studenta.

Przewaga techniki doboru parami nad randomizacją polega przede wszystkim na możliwości dokładniejszego ujednolicenia grup. Rozkłady zmiennych określonych z góry jako istotne i ewentualnie początkowy rozkład zmiennej zależnej są wyrównane przez celowy dobór w parach. Pozostałe zmienne wyrównane są przez losowy podział w parach. Dokładniejsze ujednolicenie obu grup sprawia, że metoda doboru parami pozwala — przy tej samej liczebności grupy — na dokładniejszą ocenę wpływu bodźca czy też, co na jedno wychodzi, do udowodnienia istotności wpływu potrzebuje grup mniej licznych. Metoda doboru parami daje najlepsze wyniki wtedy, gdy stosuje się ją do budowy grup złożonych z niewielkiej liczby bardzo zróżnicowanych jednostek. Takie grupy często trzeba tworzyć w eksperymentach w socjologii.

Istotne są też wady techniki doboru parami, niespotykane w technice randomizacji.

1. Im więcej zmiennych ma obejmować i im bardziej, precyzyjny ma być dobór parami, tym większej liczby jednostek op wymaga i tym więcej jednostek trzeba będzie odrzucić, jako nie mających swych bliźniaczych odpowiedników, z którymi mogłyby one tworzyć pary. Dobór parami powoduje ogromne zmniejszenie liczebności wyjściowego zbioru.

2. Wyrównywanie parami wymaga uprzedniej trafnej znajomości cech istotnych — błędy w tym zakresie, wpływają ujemnie na dokładność uzyskanych wyników.

3. Zmienne wyrównywane muszą być w jakiś sposób mierzalne a w poprzedniej technice nie muszą. Wyrównanie w parach może czynić koniecznym pretest, co by wymagało dodatkowej pracy i tworzyło tylekroć już omawiane kłopoty związane z jego niepożądanym wpływem na zmienną zależną i bodziec.

4. Wreszcie, próba wyselekcjonowanych par nie jest reprezentatywna dla zbiorowości, z której ją wybrano. Zatem, omawiana technika ujednolicania grup w eksperymencie powinna być stosowana raczej w takich eksperymentach, w których ogólność uzyskanych wyników interesuje nas mało.

Rozwinięcie techniki doboru parami polega na zwiększeniu liczby jednostek wstępnie wyrównywanych ze względu na wartości zmiennych istotnych. Ceną za to musi być zmniejszenie liczby tych zmiennych, często aż do jedności. Wówczas już całą zbiorowość można podzielić — bez straty jednostek — na kategorie o takich samych wartościach zmiennych istotnych. Kategorie te nazywa się blokami. Liczba jednostek w blokach może wielokrotnie przewyższać liczbę wariantów oddziaływania. Jednostki wewnątrz każdego bloku przydziela się losowo poszczególnym wariantom. Powstaje możliwość badania efektów manipulacji w każdym bloku oddzielnie, a przedmiotem zainteresowania stają się — obok różnic między grupami poddanymi poszczególnym rodzajom manipulacji — także różnice między blokami. Ten sposób analizy zostanie wykorzystany w schemacie bloków zrandomizowanych.

Trzecią metodą wyrównywania grup w eksperymencie jest kontrola rozkładów częstości względnych. Polega ona na dobraniu dwu lub więcej grup, które posiadają jednakowe rozkłady zmiennej zależnej, a jeśli jej pretestu się nie wykonuje — jednakowe rozkłady zmiennych podejrzanych o zależności z nią. Rzecz jasna, może się zdarzyć, że część jednostek w jednej grupie nie będzie mieć swych odpowiedników w grupach pozostałych. O tym, która z grup poddana zostanie bodźcowi lub pewnemu jego wariantowi, znowu powinna decydować randomizacja.

Metoda ta poważnie zmniejsza straty w jednostkach, dotkliwe przy doborze parami i jest łatwiejsza w stosowaniu. Jednakże ujednolicenie rozkładów zmiennych istotnych nie jest tym samym co wyrównanie kombinacji zmiennych, zapewnione w obu poprzednich metodach — mimo zachowania proporcji, ich kombinacje mogą być różne. Podobnie jak w poprzednim wypadku trzeba też wiedzieć z góry, które zmienne są istotne i winny być wyrównane, a które są nieistotne i mogą być zostawione bez kontroli — błędy tej wiedzy wpływają na rzeczywisty stopień ujednolicenia grup. Jeśli z eksperymentu wnosi się o szerszej populacji, to grupy winny być dla niej reprezentatywne, a wówczas trzeba znać także strukturę tej populacji. Wreszcie, im więcej cech mamy wyrównać w ten sposób, tym stosowanie tej metody trudniejsze — liczba tych cech wyjątkowo tylko może przekraczać kilka.

Kontrolę rozkładów częstości względnych jako metodę wyrównywania grup stosował np. Milgram w opisywanych już eksperymentach nad posłuszeństwem. Badając warunki sprzyjające i niesprzyjające posłuszeństwu, Milgram tworzył grupy poddawane wpływowi różnych warunków, wyrównawszy uprzednio w tych grupach proporcje wieku i grup zawodowych badanych. Musiał on sądzić, że inne niekontrolowane różnice między grupami są bez znaczenia.

Ostatnia z technik doboru grup do eksperymentu to technika grup równoważnych, najmniej ze wszystkich sformalizoawna i — abstrakcyjnie rzecz ujmując —

najmniej pewna, lecz zarazem lepiej niż inne dostosowana do wielu dziedzin badań socjologicznych.

Gdy projektowany eksperyment objąć ma dwie lub niewiele więcej jednostek a jednostkami tymi są naturalne zbiorowości i instytucje, wówczas losowanie kilku zbiorowości z większej ich liczby lub losowy podział paru instytucji na eksperymentalne i kontrolne nie miałby sensu. Zabiegi te dałyby wyrównanie mniej dokładne niż dobór celowy, dokonany na podstawie posiadanego przez badacza rozeznania w różnicach i podobieństwach między zbiorowościami. Zatem, eksperymentator stara się wybrać, spośród danych mu do wyboru zbiorowości, zbiorowości możliwie najbardziej do siebie podobne.

Na doborze tego rodzaju opartych było wiele wartościowych terenowych eksperymentów socjologicznych. I tak np. Morse i Reimer sprawdzali eksperymentalnie hipotezę, że stopień uczestnictwa szeregowych członków organizacji w podejmowaniu decyzji organizacyjnych ma wpływ na ich wydajność pracy i zadowolenie z pracy[77]. Eksperymentatorzy ci wybrali cztery działy pewnej instytucji, które były wyrównane ze względu na zmienne zależne a także takie zmienne istotne, jak alokacja decyzji, typ pracy, skład personelu i styl nadzoru. W dwóch z tych działów zrealizowano program zwiększenia autonomii pracowników a w dwóch zrealizowano program zmniejszenia tej autonomii. Pomiar zmiennych zależnych wykazał, że potwierdziła się tylko hipoteza o wpływie współuczestnictwa na zadowolenie z pracy; zdaniem badaczy instytucja, w której przeprowadzono eksperyment. mogła nie być odpowiednią dla sprawdzania hipotezy o wpływie współuczestnictwa na wydajność. Morse i Reimer nie informują, w jaki sposób podzielili wybrane działy na dwie grupy. Gdyby jednakże zastosowali oni podział losowy, to i tak nic by on nie dał ze względu na liczbę dzielonych jednostek.

Istnieją interesujące sposoby podwyższania trafności wewnętrznej wyników eksperymentów opartych na wyłącznie celowym doborze grup.

Jeden z nich polega na wprowadzeniu bodźca do tych zbiorowości spośród kilku zbiorowości dobranych nielosowo, w których wartości zmiennych zależnych i istotnych działają na niekorzyść sprawdzanej hipotezy. Jeżeli domniemany w hipotezie wpływ bodźca miałby polegać we wzroście wartości zmiennej zależnej, to grupą eksperymentalną winna być grupa, w której zmienna ta ma wartość niższą niż w drugiej grupie. Jeżeli po wprowadzeniu bodźca, stosunek ten się odwróci, będzie to znaczyć, że hipoteza się potwierdziła. Gdyby postąpić odwrotnie i bodziec wprowadzić do grupy o wyższej wartości zmiennych zależnych, to wyniki byłyby mniej przekonujące.

Ten sposób alokacji bodźców zastosowany został w eksperymencie Bowersa i Seashore'a z zakresu socjologii organizacji[78]. Badacze ci sprawdzali hipotezy mówiące, że sprawność funkcjonowania organizacji przemysłowej jest większa wtedy, gdy nadzór obejmuje grupy a nie jednostki, gdy interakcje w grupach są intensywne, gdy pracownicy współuczestniczą w podejmowaniu decyzji, i gdy stosunek zwierzchników' do pracowników jest życzliwy. Zmiennymi zależnymi było zadowolenie z pracy, absencja, wydajność pracy i wielkość strat w produkcji. Do eksperymentu wybrano pięć wydziałów pewnej fabryki, podobnych do siebie pod wieloma względami, przede wszystkim — stażem pracy i kwalifikacjami zatrudnionych. Nie wszystkie działy były równe ze względu na zmienne zależne. W trzech działach wprowadzono zmiany w zarządzaniu wymagane dla sprawdzenia hipotezy a dwa pozostawione bez zmian jako działy kontrolne. Jednak podziału dokonano tak, że zmianami objęto wydziały o niższych wartościach zmiennej zależnej, a wydziały o wyższych wartościach służyły do kontroli. Wykonany po pewnym czasie posttest i niezbędne porównania pokazały, że proporcje wyjściowe odwróciły się, co przekonująco wyklucza możliwość wyjaśnienia różnic wyników różnicami początkowymi i podnosi stopień potwierdzenia sprawdzanych hipotez.

[77] Morse, Reimer, op. cit.

[78] D. G. Bowers, S. E. Seashore, Changing the structure and functioning of an organization, op. cit.

Inny sposób polega na tym, że każda z dwu grup w eksperymencie występuje kolejno w roli grupy eksperymentalnej i w roli grupy kontrolnej. Wpływ nierówności między tymi grupami na uzyskane wyniki zostaje więc wyeliminowany. Sposób ten ograniczony jest do takich tylko bodźców, które mają wpływ przejściowy. Jest stosowany jako technika doboru grup w schemacie rotacyjnym i przy opisie tego schematu został on dokładnie przedstawiony.

Pewna tradycja terminologiczna odmawia miana eksperymentu takim badaniom, które posługują się technikami doboru grup nie wymagającymi randomizacji. Być może decyzja taka jest uzasadniona w naukach, w których randomizacja jest łatwo wykonalna: biologia, agrotechnika, a nawet psychologia należą do tych nauk. Nie wydaje się jednak słuszne przestać nazywać eksperymentami a zacząć nazywać quasi- -eksperymentami badań, które różnią się od bezspornych eksperymentów tylko techniką doboru grup, zwłaszcza, że dobór celowy w socjologii jest częstokroć jedyną techniką z sensem stosowalną.

*

* *

Przegląd technik wyrównywania grup i alokacji bodźców w eksperymencie posłużył nam za dopełnienie prezentacji jednej grupy schematów eksperymentalnych. Stanowi on zarazem wstęp do prezentacji drugiej grupy takich schematów.

6.3. SCHEMATY WIELOWARIANTOWE

Podejście do planowania i interpretacji eksperymentów, które przedstawimy w tym podrozdziale, cechują — wg Cochrana — następujące właściwości:

> „1. Żądanie, by sam eksperyment dostarczał znaczącego oszacowania podstawowej zmienności wyników pomiarów reakcji na warianty (odmiany) manipulacji.
>
> 2. Zastosowanie randomizacji w celu dostarczenia tych oszacowań zmienności.
>
> 3. Zastosowanie blokowania w celu zrównoważenia znanych ubocznych źródeł zmienności.
>
> 4. Zasada, że statystyczna analiza wyników wyznaczona jest przez sposób przeprowadzenia eksperymentu.
>
> 5. Koncepcja eksperymentu czynnikowego-, która kładzie nacisk na korzyści badania efektów różnych czynników lub zmiennych w pojedynczym złożonym eksperymencie, zamiast poświęcenia oddzielnego eksperymentu każdemu czynnikowi”[79].

Przedstawimy teraz najważniejsze schematy eksperymentalne oparte na tych założeniach: schemat kompletnej randomizacji, schemat bloków zrandomizowanych i schemat kwadratu łacińskiego; omówimy też zasady eksperymentu czynnikowego[80].

6.3.1. Schemat kompletnej randomizacji

Spotkaliśmy się już z takim, bardzo zresztą popularnym, schematem eksperymentalnym, który polegał na podziale zbiorowości na dwie ujednolicone grupy a następnie — poddaniu jednej z nich, działaniu bodźca, zmierzeniu zmiennej zależnej w obu grupach i obliczeniu różnicy w wynikach pomiaru i uznaniu jej za

[79] W. G. Cochran, Experimental design: The design of experiments, w: International encyclopedia of the social sciences, New York 1968, vol. 5, s. 245.

[80] Wykład schematów wielowariantowych opieramy na nast. podręcznikach: W. G. Cochran, G. M. Cox, Experimental designs, New York 1957; D. R. Cox, The planning of experiments, New York 1958; A. L. Edwards, Experimental design in psychological research, New York 1968; V. Keith, Design and analysis in experimentation, Ottawa 1972; W. Oktaba, Elementy statystyki matematycznej i metodyka doświadczalnictwa, Warszawa 1977.

miarę wpływu bodźca. Najlepszą techniką podziału zbiorowości na grupy wyrównane był podział losowy.

Pozwalał on obyć się bez pretestu. Schemat ten nazywaliśmy schematem z grupą kontrolną bez protestu.

Rozwinięciem tego schematu i techniki doboru grup jest schemat kompletnej randomizacji, zwany też schematem grup zrandomizowanych.

Schemat kompletnej randomizacji polega po prostu na losowym podziale jednostek balanych na grupy przydzielone poszczególnym wariantom, przy czym wariantów tych może być więcej niż dwa a jednym z nich może być brak oddziaływania. (Korzyści losowego podziału omówiliśmy szczegółowo w poprzednim punkcie). Do pomiaru efektów wariantów nie jest potrzebny pretest.

Oto schemat i zapis wyników eksperymentu opartego na schemacie kompletnej randomizacji. W eksperymencie tym 20 jednostek rozdzielono losowo na 4 warianty, po 5 jednostek na każdy wariant; jeden z wariantów polegał na braku oddziaływania.

		Warianty (odmiany)			
		X_0	X_1	X_2	X_3
Jednostki	1	Y_{01}	Y_{11}	Y_{21}	Y_{31}
	2	Y_{02}	Y_{12}	Y_{22}	Y_{32}
	3	Y_{03}	Y_{13}	Y_{23}	Y_{33}
	4	Y_{04}	Y_{14}	Y_{24}	Y_{34}
	5	Y_{05}	Y_{15}	Y_{25}	Y_{35}

W zapisie tym przykładowo Y_{21} oznacza wynik pomiaru zmiennej zależnej wykonanego na pierwszej z jednostek przydzielonych wariantowi *X2*.

Zaletami schematu kompletnej randomizacji są: elastyczność samego schematu oraz łatwość i prostota analizy wyników.

Schemat ten dopuszcza każdą liczbę wariantów. Liczba jednostek dla poszczególnych wariantów, liczba wariantów i liczba jednostek dla poszczególnego wariantu nie musi być równa; nawet gdy część danych dla pewnych jednostek lub wariantów trzeba z jakichś powodów odrzucić lub nie można ich uzyskać, to można prowadzić analizę na pozostałych jednostkach i wariantach.

Standardową metodą analizy wyników w tym schemacie a także w pozostałych schematach wielowariantowych jest analiza wariancji. Analiza wariancji polega na podziale całkowitej wariancji wyników na wariancje pochodzące od poszczególnych źródeł zmienności. Oparta jest na założeniu normalności rozkładów zmiennej zależnej i początkowej równości wariancji w grupach jednostek przydzielonych poszczególnym odmianom. Istotność różnic między wariancjami końcowymi tych grup mierzy się testem F. Wylicza się go jako stosunek wariancji między odmianami (zwanej też systematyczną) do wariancji błędu, powodowanej przez czynniki niekontrolowane.

Analiza wariancji zastosowana do schematu kompletnej randomizacji pozwala podzielić całkowitą wariancję eksperymentalną na wariancję między odmianami i wariancję wewnątrz odmian. Pierwsza bierze się z różnic między odmianami a druga z niejednorodności grup jednostek przydzielonych poszczególnym odmianom — gdyby wszystkie one reagowały jednakowo na pewien wariant to wariancja

wewnątrz odmian byłaby równa zeru. Zakłada się, że odmiany wpływają na uzyskane wyniki niezależnie od czynników losowych. ,

Opisaną tu procedurę planowania i analizy eksperymentu spotkać można bardzo często w raportach badawczych. Za przykład posłuży nam eksperyment z zakresu socjologii organizacji wykonany przez Trowa[81]. Badacz ten chciał ustalić w drodze laboratoryjnej symulacji procesu organizacyjnego wpływ zmian w składzie personelu na sprawność wykonywania zadań przez organizację. Stworzył on 24 małe pięcioosobowe zespoły o ujednoliconym rozkładzie zdolności umysłowych członków; zbiorowość zespołów stała się przez to jednorodną. Obmyślił 7 wariantów manipulacji rozciągających się od całkowitej stabilności składu zespołu przez cały czas badania aż do zmian tak intensywnych, że w ciągu badania przez grupę przewinąć się miały 24 osoby. To uczyniwszy, przydzielił losowo zespoły do wariantów bodźca i rozpoczął eksperyment, w którym każdy zespół miał wykonać 8 razy to samo zadanie (uczenie się było uniemożliwione), ale każdy z nich przechodził różne wielkością i kolejnością wprowadzania zmiany składu. Okazało się m. in., że sprawność zespołów nie zależała od ogólnej ilości zmian zmierzonej w liczbach bezwzględnych. Malała ona natomiast, gdy liczba osób nowych w kolejnej próbie była większa niż zmiany dotychczasowe, do których zespół już przystosował się. Wniosek ten uzyskano w drodze analizy wariancji.

Podstawowa wada schematu kompletnej randomizacji jest ewidentna. Całe zróżnicowanie między jednostkami składa się na błąd eksperymentalny i obniża istotność wyników. Stąd schemat ten może być stosowany przede wszystkim tam, gdzie spodziewane są duże różnice w efektach poszczególnych rodzajów manipulacji (rośnie wtedy wariancja systematyczna) albo tam, gdzie mamy do czynienia ze zbiorowościami bardzo jednorodnymi (maleje wtedy wariancja błędu). Gdy z jakichś powodów nie można obmyślić takich manipulacji i dobrać takich zbiorowości, pozostaje tylko zwiększanie efektywności eksperymentu przez doskonalenie schematów. Polegać ono będzie na czymś innym — na kontroli coraz większej części zmienności wyników nie pochodzących z różnic między wariantami; błąd będzie wtedy malał a istotność wyników rosła.

6.3.2. Schemat bloków kompletnie zrandomizowanych

Tak jak schemat kompletnej randomizacji stanowił uogólnienie schematu z grupą kontrolną bez pretestu, w którym grupy dobierano losowo, tak schemat bloków kompletnie zrandomizowanych stanowi rozwinięcie tego samego schematu, lecz z grupami wyrównywanymi techniką doboru parami.

Schemat bloków kompletnie zrandomizowanych stosuje się tam, gdzie jednostki stanowią zbiorowość heterogeniczną ze względu na zmienne zależne lub zmienne istotne. Schemat kompletnej randomizacji obciążałby wówczas wynik zbyt dużym błędem. Stosuje się go też tam, gdzie efekty manipulacji nie są duże i wykrycie ich w eksperymencie opartym na schemacie poprzednim wymagałoby zbyt dużej liczby jednostek.

Schemat ten wymaga wstępnego pogrupowania jednostek w jednorodne bloki. Blokiem może być grupa osób tworzących pewną zbiorowość naturalną, np. klasa szkolna lub wyodrębniona przez operacje wyrównujące, np. osoby o tej samej płci. Blokiem może być para jednostek, a nawet jedna jednostka, jeśli schemat eksperymentu przewiduje poddawanie jej kolejno poszczególnym wariantom manipulacji. Tworząc blok, zakłada się, że reakcje w bloku będą mniej zróżnicowane niż w całej zbiorowości. Wartość schematu bloków kompletnie zrandomizowanych zależy od trafności tego założenia. Tworzenie bloków na podstawie zmiennych nieistotnych obniża dokładność eksperymentu.

Drugi etap w eksperymentowaniu opartym na omawianym schemacie to losowy przydział jednostek w bloku poszczególnym wariantom. Randomizację wykonuje się w każdym bloku oddzielnie. W ten sposób każdy blok jest powtórzeniem eksperymentu.

[81] D. B. Trow, Membership succession and team performance, w: W. M. Evan (ed.), Organizational experiments, New York 1971.

Oto zapis schematu bloków kompletnie zrandomizowanych dla 5 bloków i 4 wariantów.

		Warianty (odmiany)			
		X_0	X_1	X_2	X_3
Bloki	I	1	3	2	4
	II	4	2	3	1
	III	1	4	2	3
	IV	2	3	1	4
	V	3	1	4	2

Zauważmy, że liczba bloków nie musi być równa liczbie wariantów zwykle ta pierwsza przewyższa drugą.

Analiza wyników eksperymentu opartego na schemacie bloków kompletnie zrandomizowanych polega na podziale wariancji całkowitej na wariancję między odmianami, między blokami i wariancję błędu. Zakłada się niezależność wymienionych źródeł zmienności. Test *F* pozwala ocenić istotność różnic między odmianami a także między blokami.

Łatwo teraz pokazać różnice w dokładności wyników między schematem kompletnej randomizacji a schematem bloków kompletnie zrandomizowanych. W poprzednim schemacie wariancję wewnątrz odmian uznaje się za błąd, w obecnym zaś błąd ten zmniejsza się o wariancję międzyblokową. To jest największa zaleta tego schematu. Analiza w blokach ma jeszcze jedną zaletę. Gdy różnice między wariantami we wszystkich blokach są takie same, to uzyskane wyniki zyskują na pewności. Gdy są różne, to zyskujemy wiedzę o modyfikatorach ustalonych zależności. Gdyby eksperyment, który dał różne wyniki w poszczególnych - blokach, oparty był na schemacie kompletnej randomizacji, to efekt manipulacji nie byłby widoczny.

Wadą tego schematu w stosunku do poprzedniego jest konieczność uzupełniania braków pewnych obserwacji przez dodatkowe obserwacje lub dodatkowe rachunki.

Z dokonanego porównania wynika, że schemat bloków kompletnie zrandomizowanych jest schematem bardziej wartościowym niż schemat kompletnej randomizacji. Randomizacja zapewnia jedynie, że każdy wariant ma równą szansę bycia faworyzowanym lub upośledzonym przez zewnętrzne źródła zmienności. Blokowanie — efekty tych źródeł eliminuje. Stąd blokowanie winno być używane do. kontroli ważnych znanych zakłóceń, a randomizacja do kontroli zakłóceń mniej ważnych. Schemat bloków kompletnie zrandomizowanych łączy obie te metody.

6.3.3. Schemat kwadratu łacińskiego

Schemat kompletnej randomizacji całą zmienność nie pochodzącą od wariantów zaliczał do błędu. Schemat bloków zrandomizowanych pozwalał kontrolować zmienność pochodzącą od pewnej zmiennej ubocznej i zmniejszał tym samym wariancję błędu. Jeszcze dalej idzie schemat zwany kwadratem łacińskim — ten umożliwia kontrolę zmienności pochodzącej od drugiej zmiennej ubocznej.

Kwadrat łaciński można zastosować wtedy, gdy każda ze zmiennych przyjmuje tyle samo wartości i liczba tych wartości równa się liczbie wariantów bodźca.

Powiedzmy, że zamierzamy ocenić eksperymentalnie wpływ trzech systemów wynagradzania za pracę na wydajność pracy zespołowej. Wiemy zarazem, że wy-

dajność zależy od kwalifikacji formalnych pracowników i że badane brygady zróżnicowane są pod tym względem. Powinniśmy wówczas stworzyć 3 bloki o wyrównanych poziomach kwalifikacji i w każdym z nich losowo przydzielić brygady do poszczególnych systemów wynagradzania. Ale wiemy też, że wydajność pracy zespołowej zależy od stopnia integracji mierzonej liczbą lat przepracowanych razem przez członków brygady i że brygady różnią się integracją. Gdybyśmy w tej sytuacji poprzestali na układzie bloków kompletnie zrandomizowanych, zróżnicowanie integracji spowodowałoby duży błąd eksperymentalny. Moglibyśmy więc stworzyć 9 grup 3-brygadowych, każda o innej kombinacji obu zmiennych ubocznych — kwalifikacji i integracji (obie one musiałyby mieć po trzy wartości) a następnie brygady w każdej z tych grup przydzielić losowo poszczególnym wariantom. Możemy jednakże obejść się bez zwiększenia liczby brygad. Wystarczy tak kombinować wartości zmiennych zewnętrznych i tak podzielić warianty na kombinacje, by każdy określony wariant bodźca pojawiał się za każdym razem przy innej kombinacji i by przy każdej kombinacji działał tylko jeden wariant.

Wymagania te spełnia schemat kwadratu łacińskiego. W schemacie tym bloki jednego rodzaju nazywają się wierszami a drugiego rodzaju kolumnami. Dla eksperymentu, w którym liczba bloków równa liczbie wierszy i równa liczbie wariantów wynosi 4, zapis schematu zamieszczamy poniżej; literami oznaczono poszczególne warianty (odmiany) bodźca.

		Kolumny			
		K_1	K_2	K_3	K_4
Wiersze	W_1	*B*	*C*	*D*	*A*
	W_2	*C*	*D*	*A*	*B*
	W_3	*D*	*A*	*B*	*C*
	W_4	*A*	*B*	*C*	*D*

W schemacie tym widać, że każdy wariant pojawia się raz i tylko raz w każdej kolumnie i raz i tylko raz w każdym wierszu. Jest bardzo wiele kwadratów spełniających ten warunek. Randomizacja w Schemacie kwadratu łacińskiego polega właśnie na losowym wyborze jednego z kwadratów możliwych dla danej liczby wariantów.

Profit z takiego uporządkowania jest bardzo duży. Można eliminować zarówno wpływy wierszowe jak i wpływy kolumnowe, tj. wpływ podwójnego zróżnicowania. Wariancja całkowita dzieli się na wariancję międzywariantową, wariancję między wierszami, wariancję między kolumnami i wariancję błędu. Ta ostatnia zmniejszona jest w stosunku do schematu bloków zrandomizowanych o wariancję między blokami drugiego rodzaju. Zatem, udało się zwiększyć efektywność doświadczenia.

Schemat kwadratu łacińskiego zakłada[1] brak interakcji między wariantami, wierszami, kolumnami i błędem: wpływy jednych nie są zależne od drugich. Jeśli założenie to faktycznie spełnione nie jest, zwiększa się błąd. Gdy efektów wierszowych i kolumnowych faktycznie nie ma, to schemat ten jest mniej efektywny od schematu bloków zrandomizowanych; tak samo jak schemat bloków zrandomizowanych był mniej dokładny od schematu kompletnej randomizacji, gdy nie było efektów blokowych.

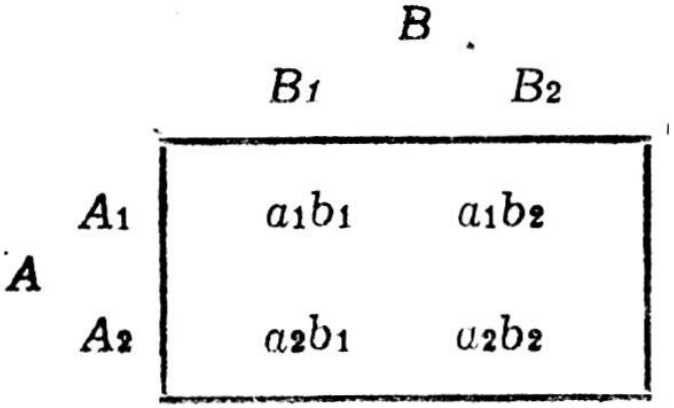
Warunki pozwalające zastosować schemat kwadratu łacińskiego (liczba wierszy równa liczbie kolumn, równa liczbie wariantów) mogą być trudne do spełnienia. Łatwiej je spełnić wtedy, gdy zmienne blokowane nie są cechami naturalnymi, takimi jak płeć czy wykształcenie, lecz mogą być kształtowane przez badacza, np. poziom motywacji czy lęku. Łatwiej spełnić je również wtedy, gdy zmienne te mają charakter ciągły. Można wówczas wybrać taką liczbę przedziałów zmiennych, jaka jest wymagana przez schemat eksperymentu..

6.3.4. Schematy czynnikowe

Ostatnie z omawianych schematów eksperymentalnych pozwalały na badanie wpływu zmiennych przybierających wiele wartości. Nie pozwalały jednak na badanie wpływu wielu zmiennych. Ograniczenie to pokonują schematy czynnikowe — w eksperymentach na nich opartych można badać równocześnie wpływ wielu zmiennych, z których każda może przybierać wiele wartości.

W schematach czynnikowych badane zmienne nazywa się czynnikami a ich wartości — poziomami. W schematach poprzednich rolę wariantów manipulacji grały poziomy danego czynnika, w schematach czynnikowych rolę tę przejmują kombinacje poziomów różnych czynników. Liczba kombinacji możliwych jest iloczynem liczb poziomów poszczególnych czynników. Jeśli liczba wariantów zrealizowanych w pewnym eksperymencie jest równa temu iloczynowi, eksperyment ten nazywamy kompletnym eksperymentem czynnikowym. Jeśli w dodatku każdemu z wariantów poddano tę samą liczbę jednostek, eksperyment nazywamy kompletnym eksperymentem czynnikowym z równą liczbą replikacji. Istnieją też schematy niekompletne a także schematy o różnej liczbie replikacji. My mówić będziemy jednak tylko o schematach kompletnych z równą liczbą replikacji: Eksperyment czynnikowy może być oparty na każdym z omówionych wyżej schematów wielowariantowych: schemacie kompletnej randomizacji, schemacie bloków zrandomizowanych i schemacie kwadratu łacińskiego.

Najprostszy schemat czynnikowy wymaga dwóch czynników, każdy o dwóch poziomach; oznacza się go iloczynem 2×2 lub potęgą 2^2, Schemat 2×2 posłuży nam za model schematów czynnikowych, na którym pokażemy zasady analizy wyników eksperymentów na nich opartych. Tabela poniższa przedstawia wyniki eksperymentu czynnikowego opartego na schemacie 2×2. Literami *A* i *B* oznaczono czynniki, literami *A* i *B* z askryptami — poziomy czynników, a literami *ab* z odpowiednimi askryptami oznaczono średnie wyniki pomiaru zmiennej zależnej dla każdego z wariantów. Zakładamy, że liczba jednostek poddanych każdemu z wariantów jest ta sama. Jest to więc kompletny eksperyment czynnikowy z równą liczbą replikacji.



Schemat ten pozwala wyliczyć efekty każdego z czynników oddzielnie. Rozróżnia się efekty proste i efekt główny danego czynnika. Efektem prostym pewnego czynnika jest zmiana wartości zmiennej zależnej zaobserwowana przy zmianie poziomu tego czynnika a przy stałym poziomie pozostałych czynników. Tak więc,

efektem prostym czynnika *A* przy poziomie *B1* jest różnica *a2b1 - a1b1,* a przy poziomie *B2* różnica *a2b2— a1b2*.

Efektem głównym nazywa się średnią efektów prostych. Tak więc, efektem głównym czynnika *A* jest wielkość równa

$$\frac{1}{2}[(a_2b_1 - a_1b_1) + (a_2b_2 - a_1b_2)].$$

W ten sam sposób można wyliczyć efekty proste i efekt główny czynnika *B*.

Wartość schematu czynnikowego ujawnia się wtedy, gdy przechodzimy do badania połączonego wpływu czynników.

Efekty proste mogą być sobie równe i tym samym, mogą być równe efektowi głównemu. Znaczy to tyle, że wpływ jednego czynnika nie zależy od czynników pozostałych. Zazwyczaj jednak tak nie jest: efekt pewnego czynnika zmienia się zależnie od poziomu innych czynników. Schemat czynnikowy pozwala mierzyć takie interakcje. Miary interakcji są funkcją różnic między efektami prostymi. W schemacie 2 X 2 interakcja wyraża się wielkością równą

$$\frac{1}{2}[(a_2b_1 - a_1b_1) - (a_2b_2 - a_1b_2)].$$

Sposób rozwijania schematu 2×2, najprostszego ze schematów czynnikowych, łatwo zasugerować na schemacie obejmującym 3 czynniki, każdy na 2 poziomach; schemat ten oznacza się iloczynem 2×2×2 lub potęgą 2^3. Oto zapis takiego schematu:

	C_1		C_2	
	A_1	A_2	A_1	A_2
B_1	$a_1b_1c_1$	$a_2b_1c_1$	$a_1b_1c_2$	$a_2b_1c_2$
B_2	$a_1b_2c_1$	$a_2b_2c_1$	$a_1b_2c_2$	$a_2b_2c_2$

W schemacie 2’ istnieje 8 wariantów. Każdy z czynników ma 4 efekty proste — dla każdej z kombinacji poziomów pozostałych dwu czynników. Efekt główny , danego czynnika jest przeciętną tych 4 efektów prostych.

W schemacie 2^3 występują dwa rodzaje interakcji: interakcje między dwoma czynnikami, zwane interakcjami pierwszego rzędu i interakcje między trzema czynnikami, zwane interakcjami drugiego rzędu. Aby je wyliczyć, trzeba najpierw wyliczyć interakcje między dwoma, czynnikami dla każdego poziomu trzeciego czynnika oddzielnie. Tak więc, interakcja *A* × *B* dla *C1* równa się

$$\frac{1}{2}[(a_2b_1c_1 - a_1b_1c_1) - (a_2b_2c_1 - a_1b_2c_1)]$$

Natomiast interakcja A × B dla C2 równa się

$$\frac{1}{2}\left[(a_2b_1c_2 - a_1b_1c_2) - (a_2b_2c_2 - a_1b_2c_2)\right]$$

Średnia tych interakcji stanowi interakcję $A \times B$ pierwszego rzędu. Różnica tych ' interakcji dzielona przez 2 jest miarą interakcji $A \times B \times C$ drugiego rzędu. Wielkość ta mówi nam, czy interakcja między dwoma czynnikami jest taka sama dla każdego poziomu trzeciego czynnika. Jeśli jest ona równa O, znaczy to, że wpływ poziomu jednego czynnika na wpływ drugiego czynnika na zmienną zależną jest niezależny od poziomu trzeciego czynnika. Interakcje drugiego rzędu rzadko się bada: są one małe w porównaniu z efektami głównymi i interakcjami dwuczynnikowymi i nieistotnie różnią się od błędu.

Metodą analizy wyników eksperymentów opartych na schematach czynnikowych jest zazwyczaj **analiza** wariancji. Polega ona tutaj na podziale całkowitej wariancji na wariancję pochodzącą od poszczególnych czynników, od interakcji pierwszego rzędu, od interakcji drugiego rzędu i na wariancję błędu. Jeśli eksperyment założono na planie bloków zrandomizowanych, wyodrębnia się dodatkowo wariancję międzyblokową, a jeśli założono go na schemacie kwadratu łacińskiego — wariancję wierszową i kolumnową.

Zwróćmy uwagę, że w schematach dotąd omawianych zakładało się brak interakcji między poszczególnymi źródłami wariancji i jeśli założenie to nie było prawdziwe, interakcja powiększała błąd eksperymentalny. W eksperymencie czynnikowym nie tylko, że nie czyni się takich założeń, lecz interakcje te mierzy się. Zwróćmy też uwagę, że w schematach dotąd omawianych blokowanie jednostek wedle wartości zmiennych istotnych było tylko zabiegiem mającym uczynić wyraźnym wpływ zmiennej eksperymentalnej. W schematach czynnikowych każda zmienna staje się samoistnym przedmiotem zainteresowania, uważana. jest za zmienną eksperymentalną.

Instruktywnym przykładem eksperymentu czynnikowego jest badanie Daya i Hamblina nad następstwami różnych stylów nadzoru pracowników[82].

Autorzy ci postawili hipotezę, że nadzór ścisły oraz nadzór punitywny będą wywoływać u pracowników frustrację przeradzającą się w agresję wobec przełożonych i obniżenie wydajności pracy. Przewidywali oni także, że zależność ta będzie silniejsza w grupie pracowników o niskiej samoocenie, ponieważ ścisły i punitywny nadzór sugerują brak kompetencji lub motywacji i stanowią w ten sposób zagrożenie dla chwiejnej samooceny.

W celu sprawdzenia swych hipotez Day i Hamblin utworzyli dwudziestoczteroosobowe zespoły o wyrównanych zmiennych istotnych i rozlosowali je po 6 na każdy z czterech wariantów nadzoru (nadzór ścisły punitywny, ścisły niepunitywny, ogólny punitywny i ogólny niepunitywny). Zespołom tym polecili wykonywać pewne zadanie w warunkach upozorowanych na fabryczne; Był to kompletny eksperyment czynnikowy 2×2 założony na schemacie kompletnej randomizacji. Eksperymentatorzy dokonali też pomiaru samooceny badanych i wprowadzili ją jako trzeci czynnik do analizy wyników. W ten sposób eksperyment przyjął postać eksperymentu 2 × 2 × 2, w którym jeden z czynników był niemanipulowanym.

Analiza wariancji uzyskanych wyników wykazała, że na poziomie jednostek wystąpiły efekty główne ścisłości i punitywności nadzoru zarówno dla agresji jak i dla wydajności pracy; efekt główny samooceny nie wystąpił. Analiza wykazała również, że z 8 możliwych interakcji (ścisłość nadzoru × punitywność, ścisłość × samoocena, punitywność × samoocena, ścisłość × punitywność × samoocena; każda z interakcji dla dwu zmiennych zależnych), istotne były tylko dwie: interakcja, ścisłości nadzoru z karaniem w obniżaniu wydajności i ścisłości z samooceną w zwiększaniu agresji. Omówimy je dokładniej. Tab. 1 pokazuje pierwszą z interakcji a tab. 2 — drugą.

[82] R. C. Day, R. L. Hamblin, Effects of close and punitive styles of supervision, op. cit.

Tabela 1

		Nadzór	
		punitywny	niepunitywny
Nadzór	ogólny	100	77
	ścisły	75	67

Liczby w tabeli określają wydajność pracy w stosunku do jednostek pracujących w warunkach najbardziej sprzyjających wydajnej pracy. Widać, że aczkolwiek każdy z czynników wpływa na obniżenie wydajności pracy, to podwójna frustracja wcale nie prowadzi do podwójnego obniżenia wydajności pracy: zamiast spadku wydajności o 48% w kategorii osób poddanych nadzorowi ścisłemu i punitywnemu, wydajność ta spadła tylko o 33%; połowa tej różnicy między tymi wskaźnikami jest miarą interakcji (można ją też wyliczyć w sposób bardziej metodyczny — przez porównanie efektów prostych). Mamy tu do czynienia z sytuacją, w której obecność jednego czynnika zmniejsza efekt drugiego czynnika; ten rodzaj interakcji nazywa się niekiedy interakcją antagonistyczną.

Tabela 2

		Nadzór	
		ogólny	ścisły
Samoocena	wysoka	1,64	1,66
	niska	1,54	1,74'

Spojrzenie na wyniki uzyskane na skali agresji wobec przełożonych przez osoby przydzielone poszczególnym wariantom nadzoru przekonuje, że ścisły nadzór zwiększa agresję tylko wśród osób o niskiej samoocenie; efekt prosty ścisłości nadzoru jest nieistotny w grupie osób o wysokiej samoocenie. Efekt główny ścisłości nadzoru pochodził z tego właśnie efektu prostego i musiał być słabszy niż stwierdzony efekt prosty. W analizowanej sytuacji obecność jednego czynnika zwiększa efekt drugiego; ten rodzaj interakcji nazywa się katalitycznym.

Eksperymenty w których bada się równocześnie wpływ wielu czynników posiadają istotne przewagi nad eksperymentami, w których bada się kolejno wpływ pojedynczych czynników.

1. Gdy czynniki działają niezależnie, efekty główne każdego z czynników oszacowane są z bardzo dużą dokładnością a przy tym stosunkowo niewielkim kosztem. „Gdy wszystkie czynniki są pod względem swych efektów niezależne — piszą Cochran i Cox — ujęcie czynnikowe będzie dawać w wyniku znaczną oszczędność czasu i materiału przeznaczonego na eksperymenty. Oszczędność ta wynika z dwóch faktów. Po pierwsze, […] gdy czynniki są niezależne, wszystkie efekty proste pewnego czynnika są równe jego efektowi głównemu, tak że efekty główne są jedynymi wielkościami potrzebnymi do pełnego opisania konsekwencji zmienności danego czynnika. Po drugie, w eksperymencie czynnikowym każdy z efektów głównych oszacowany jest z taką samą dokładnością, jak gdyby cały eksperyment poświęcony był tylko temu jednemu czynnikowi. […] Gdyby było n czynników, wszystkie o dwóch poziomach 'i wszystkie niezależne, ujęcie pojedynczo-czynnikowe wymagałoby *n* razy większego materiału eksperymentalnego niż ujęcie czynnikowe o tej samej dokładności”[83].

2. Ujęcie czynikowe pozwala na badanie interakcji między czynnikami. „Gdy czynniki nie są niezależne — kontynuują Cochran i Cox — efekty proste pewnego czynnika zmieniają się stosownie do poszczególnych kombinacji innych czynników, z którymi one łącznie występują. W tym przypadku ujęcie pojedynczo-czynnikowe może dostarczać tylko pewnej liczby niepowiązanych fragmentów informacji, które niełatwo ująć w jedną całość. Aby przeprowadzić pewien eksperyment na

[83] Cochran, Cox, op. cit., s. 150 - 1.

pewnym pojedynczym czynniku *A*, musi się podjąć decyzję co do poziomów innych czynników, powiedzmy *B*, C, *D*, które mają być użyte w tym eksperymencie (…) Eksperyment ujawnia efekty *A* dla tej szczególnej kombinacji *B, C* i *D,* ale nie dostarcza żadnych informacji umożliwiających przewidywanie efektów *A* dla jakiejś innej kombinacji czynników *B, C* i *D*. Z drugiej strony, w ujęciu czynnikowym, efekty *A* zbadane są dla każdej kombinacji *B, C* i *D* objętej eksperymentem. W ten sposób, zostaje zgromadzona duża ilość informacji zarówno o efektach czynników jak i o ich wzajemnych oddziaływaniach”[84],

3. Wyniki eksperymentu czynnikowego mają dużą ogólność, gdyż uwzględniają rozmaite modyfikatory zależności między zmiennymi, które mogą się pojawić w sytuacjach pozaeksperymentalnych. Gdyby eksperyment skupiony był na pojedynczym czynniku, jego wyniki nie miałyby tak dużej ogólności. Stąd, rekomendacje praktyczne oparte na eksperymentach czynnikowych są bardziej wartościowe niż oparte na eksperymentach pojedynczo-czynnikowych.

Z przytoczonej listy zalet nie należy wnosić o konieczności planowania wielkich eksperymentów czynnikowych, w których badało by się równoczesny wpływ wielu wielopoziomowych czynników. Aczkolwiek nie ma żadnej teoretycznej granicy dla mnożenia czynników i poziomów, istotne względy metodologiczne nakazują wstrzemięźliwość w tym względzie.

Po pierwsze, prowadzenie takich eksperymentów jest kłopotliwe. Im więcej czynników i poziomów eksperyment miałby obejmować, tym trudniej go zorganizować. Duży eksperyment czynnikowy wymaga dużej liczby jednostek i dużej liczby oddziaływań, instrukcji i pomiarów — mnożenie i komplikacja procedur wpływa ujemnie na zachowanie badanych i rozmywa efekty czynników. Co gorsza, im większy eksperyment, tym trudniej w sposób zrozumiały interpretować jego wyniki.

Po wtóre, częstokroć duży eksperyment jest niepotrzebny. Na początku serii badań zaleca się schematy prostsze dobre do uchwycenia najważniejszych czynników; w dalszym ciągu badań schematy te można komplikować, wiadomo już jednak, jakie zmienne wziąć pod uwagę. Gdy posiada się już pewne informacje o czynnikach istotnych, zamiast planować duży kompletny eksperyment czynnikowy, lepiej skupić się na intensywnym badaniu wybranych czynników lub wybranych kombinacji ich poziomów, w celu znalezienia np. kombinacji najbardziej efektywnych. Wreszcie, choć złożone eksperymenty czynnikowe są nieocenione do sporządzenia opisu zachowania się systemu w różnych warunkach, to wyjaśnienie tego zachowania wymagać może skupienia się na kilku zmiennych podstawkowych, stąd lepsze będą wówczas eksperymenty prostsze.

Jest rzeczą oczywistą, że nie można sformułować żadnej uniwersalnej a zarazem precyzyjnej reguły, która by określała, jak złożony winien być eksperyment. Można powiedzieć tylko tyle, że „powinien on być złożony tylko w takim stopniu, jaki jest konieczny dla wyraźnego wystąpienia ważnych zależności”[85].

Na tym kończymy przegląd schematów eksperymentalnych. Ich lista nie jest zamknięta w dwojakim sensie. Po pierwsze, objęła ona tylko schematy najczęściej spotykane w badaniach społecznych. Po drugie, pomysłowość badaczy wsparta na rozwoju narzędzi pomiaru i analizy danych sprawia, że ciągle pojawiają się schematy nowe, lepiej od dotychczasowych pozwalające badaczom radzić sobie z niedostępnością i komplikacjami materii społecznej.

[84] Ibidem, s. 151.

[85] E. Aronson, J.M Carlsmith, Experimentation in social psychology, w: G. Lindzey, E. Aronson (eds.), The handbook of social psychology, Reading, Mass. 1968. vol. 2, s. 51.

7. EKSPERYMENTY SPOŁECZNE

Jedną z najważniejszych, a zdaniem niektórych poważnych obserwatorów najważniejszą, tendencją rozwojową w metodologii badań społecznych w ostatnich paru latach jest rozwój metod badań ewaluacyjnych. Badania ewaluacyjne „ustalają […] skutki działań politycznych i programów dla ich adresatów (jednostek, grup, instytucji, krajów) w terminach celów, które miały one osiągnąć. Przy pomocy obiektywnych i systematycznych metod, badania ewaluacyjne pozwalają określić stopień, w jakim cele te zostały osiągnięte i obserwować czynniki związane z sukcesami i niepowodzeniami. Zakłada się, że poprzez dostarczenie «faktów», ewaluacja pomaga decydentom w dokonywaniu mądrych wyborów między przyszłymi sposobami postępowania. Staranne i niestronnicze dane dotyczące konsekwencji programów powinny udoskonalić podejmowanie decyzji”[86]. Dwiema podstawowymi metodami badań ewaluacyjnych są analizy reform społecznych oraz eksperymentalne próby* zmian społecznych. Choć eksperymenty z wielu łatwych do uchwycenia powodów są mniej liczne niż studia nad’ reformami, to jednak z wielu innych powodów — będzie o nich mowa — stanowią wzór badań ewaluacyjnych.

W niniejszym rozdziale przedstawimy eksperymentalną metodą oceny zmian społecznych. Omówimy kolejno samo pojęcie eksperymentu społecznego, charakter hipotez sprawdzanych w tego rodzaju badaniach, problemy wprowadzania, kontroli, pomiaru i oceny skuteczności eksperymentalnej innowacji społecznej oraz pożytki, jakie eksperyment społeczny dać może praktyce i nauce, wskażemy też na trudności prowadzenia eksperymentów społecznych.

7.1. POJĘCIE EKSPERYMENTU SPOŁECZNEGO

Innowacjami społecznymi nazywa się najzwyczajniej praktyki postrzegane jako nowe w pewnym systemie społecznym. Kiedy innowacje rozprzestrzenią Się na istotną część systemu, mówi się, iż dokonała się zmiana społeczna. Tak wprowadzenie innowacji, jak i zmian społecznych może być planowane i kierowane. Jeśli decydent przed podjęciem decyzji o wprowadzeniu czy zaniechaniu zmiany, wprowadzi w życie w skali ograniczonej tytułem próby innowację, a przy tym kontroluje jej działanie, to dokonuje on eksperymentu społecznego. Innymi słowy, eksperyment społeczny jest pilotażem zmiany społecznej.

Zwróćmy uwagę na fakt, że celem nadrzędnym eksperymentalnego wprowadzenia innowacji społecznej jest poznanie — ocena nowej praktyki i wypracowanie trafnej metody zmiany w szerszej skali. Zatem, sukces eksperymentu społecznego to sprawdzenie wartości nowej praktyki w sposób zgodny z regułami postępowania badawczego, nawet jeśli wynik okaże się dla tej praktyki niepomyślny. Moment ten odróżnia eksperymentalną zmianę społeczną od zmiany społecznej po prostu. Brak oczekiwanych następstw zmiany oznacza jej niepowodzenie, bo celem jej była poprawa stanu rzeczy.

Zwróćmy też uwagę, że eksperymentator nie jest przekonany o wartości innowacji; gdyby był, to by nie eksperymentował. Reformator natomiast jest przekonany o wartości reformy — gdyby nie był, to by jej nie wprowadzał.

W konkretnych przypadkach nie zawsze łatwo jest powiedzieć, z czym mamy do czynienia: z badaniem przez działanie czy tylko z działaniem. Czy sądy robotni-

[86] C. H. Weiss, Evaluation research in the political context, w: E. L. Struening, M. Guttentag (eds.), Handbook of evaluation research, Beverly Hills 1975, vol. 1, s. 13.

cze[87], tzw. eksperyment nowosądecki[88] i tzw. eksperyment w zakładach „Era”[89] były jedynie nowatorskimi próbami usprawnienia odpowiednich wycinków życia społecznego czy też decydenci wyższego niż przedsiębiorstwo i powiat szczebla, zezwalając na funkcjonowanie wybranych jednostek na zasadach innych niż powszechnie w Polsce obowiązujące, kierowali się intencją sprawdzenia skuteczności działania opartego na zasadach szczególnych? Być może, intencje te nie były im obce.

Zakres praktyk, które można obejmować testami eksperymentalnymi jest bardzo szeroki. Mieszczą się w nim struktury organizacyjne i metody zarządzania, projekty ustawodawcze, praktyki resocjalizacyjne i metody rehabilitacji, programy szkolne i metody dydaktyczne itd. Można powiedzieć, że zakres ten obejmuje metody polityki społecznej *sensu largissimo;* nie znaczy to jednak, że wszystkie praktyki w obrębie wyróżnionych dziedzin są podatne na testy eksperymentalne.

7.2. HIPOTEZY SPRAWDZANE W EKSPERYMENTACH SPOŁECZNYCH

Eksperyment społeczny, tak jak każde badanie stanowi reakcję na pewną sytuację problemową. W tym przypadku sytuację tę można scharakteryzować następująco[90]:

a. istniejący stan rzeczy jest oceniany wedle pewnych kryteriów jako niezadowalający;

b. istnieje zgoda co do tego, jak powinien wyglądać zmieniony stan rzeczy;

c. przewidywany koszt zmiany istniejącego stanu rzeczy na pożądany jest mniejszy niż koszt tolerancji *status quo;*

d. nie wiadomo, jakimi metodami trzeba przeprowadzać zmianę; wiadomo tylko, że praktyki istniejące do tego się nie nadają.

Niewiedza, o której mowa w punkcie ostatnim przybiera dwojaką postać. Może być tak, że hipotezy mówiące, iż określone działania doprowadzą do osiągnięcia pożądanego stanu rzeczy są słabo uzasadnione bądź wręcz nie istnieją, gdyż albo teoria danej dziedziny jest słabo rozwinięta albo przypadek zbyt zindywidualizowany, by teorię tę można było zastosować z wysoką pewnością wyniku. Sytuacja druga polega na równoległym funkcjonowaniu wśród praktyków i badaczy rozmaitych konkurencyjnych rozwiązań, za którymi stoją różne teorie, doświadczenia lub interesy.

Zarysowana sytuacja decyzyjna jest sytuacją niepewności, z którą związane jest ryzyko. Ideałem jest sytuacja pewności co do następstw rozważanych działań. Sposobem zbliżenia się do niej jest właśnie eksperyment społeczny. Daje on możliwość odpowiedzi na pytanie o faktyczne skutki wchodzących w grę alteranatyw.

Eksperyment jest badaniem sprawdzającym hipotezy. Hipotezy sprawdzane przez eksperymenty społeczne głoszą, że określone praktyki, działania, programy, struktury wywołują pewne skutki nieobojętne z punktu widzenia wartości. To było bardzo ogólnie powiedziane. Czasami hipotezy te mówią, że praktyki proponowane są lepsze od istniejących, czasami, że jedne propozycje są lepsze od drugich.

[87] Zob. E. Modliński, O społecznych sądach robotniczych, „Państwo i Prawo” 1962, nr 1; A. Podgórecki, Zarys socjologii prawa, Warszawa 1971, s. 304- 328.

[88] Zob. J. Pieczkowski, Nowosądecki eksperyment ekonomiczny, w: Z. T. Wierzbicki (red.), Aktywizacja i rozwój społeczności lokalnych, Wrocław 1973; J. J. Wiatr, Experiments in self-government; the Polish case, w: J. Laponce, P. Smoker (eds.), Experimentation and simulation in political science, Toronto 1972.

[89] Zob. S. Burchardt, Stosunki społeczne w przedsiębiorstwie. Studium eksperymentu, Warszawa 1974.

[90] Zob. A. Podgórecki, Charakterystyka nauk praktycznych, Warszawa 1962, s. 66 - 83.

Niekiedy chodzi o zbadanie warunków, od których zależy skuteczność pewnego programu, o wykrycie trudności jego wdrażania lub o znalezienie jego skutków ubocznych. Hipotezy te niejednokrotnie formułowane są w stylizacji wartościującej. Kryteria wartościowania w nich uwikłane zależne są od dziedziny eksperymentowania: skuteczność prawa, efektywność zarządzania, trwałość resocjalizacji, szybkość nabywania nowych umiejętności są niektórymi z nich. Wartościowanie wedle takich kryteriów nie czyni hipotez niewrażliwymi na testy empiryczne, byleby tylko kryteria te były jasno sprecyzowane. Jako oceny celowościowe, są one podatne na te testy na równi ze zdaniami czysto opisowymi.

7.3. WPROWADZANIE INNOWACJI EKSPERYMENTALNYCH W ŻYCIE

Manipulacja eksperymentalna polega na umyślnym wytworzeniu zmian, którym sprawdzana hipoteza przypisuje pewne skutki. Manipulacja w eksperymentach społecznych przybiera postać wprowadzenia sprawdzanych praktyk w życie wybranych, już istniejących systemów społecznych, albo tworzenia systemów nowych, opartych na zasadach, które chce się ocenić. Wprowadzenie w wybranych pilotowych zakładach przemysłowych w r. 1970 bodźców materialnego zainteresowania w celu sprawdzenia ich działania przed wprowadzeniem ich w całym przemyśle jest przykładem manipulacji pierwszego rodzaju. Utworzenie pod koniec lat 50-tych eksperymentalnego zakładu resocjalizacji młodocianych przestępców w Szczypiornie ilustruje drugi rodzaj manipulacji. W eksperymentach pedagogicznych obecne są oba rodzaje zmiany. Szkoły eksperymentalne są czasem szkołami zorganizowanymi na nowych zasadach a czasem szkołami, w których wybrane klasy traktuje się inaczej niż pozostałe.

Wprowadzanie eksperymentalnych innowacji w życie społeczne nasuwa pewne problemy metodologiczne; pominiemy tu banalną kwestię braku władzy do tego potrzebnej — wtedy po prostu nie ma eksperymentu.

W metodologii eksperymentowania znany jest konflikt między dążeniem do uzyskania wyraźnych różnic w wartościach zmiennej zależnej a dążeniem do kontroli wpływu manipulacji eksperymentalnej. Manipulacje złożone wywierają znaczny wpływ, lecz z drugiej strony trudno wtedy wskazać ten ich aspekt, który za wpływ ów jest odpowiedzialny. Dylemat ten obecny jest i w metodologii eksperymentów społecznych. Przed eksperymentatorem staje i tutaj pytanie: Jaka składowa zmiennej odróżniającej grupę eksperymentalną od kontrolnej sprawia, że jednostki w tych grupach zachowują się odmiennie. Pewną metodę znajdowania odpowiedzi na to pytanie bez upraszczania programu eksperymentalnego zaproponował Kiesler[91] Radzi on stosować obok grupy eksperymentalnej i kontrolnej trzecią grupę, w której wprowadzono by w życie pewną tylko część złożonego programu, któremu poddano grupę eksperymentalną. Stosownie do wielkości różnic między zmianą zaobserwowaną w grupie eksperymentalnej a zmianą w trzeciej grupie, można będzie wypowiadać się o istotności i nieistotności kontrolowanego aspektu programu eksperymentalnego dla wywoływania tej zmiany; np. może okazać się, że cała reszta tego programu jest niepotrzebna, albo że koszt komplikacji programu przewyższa wartość uzyskanej w jej wyniku zmiany.

Inny problem eksperymentowania to problem skuteczności samego wprowadzenia programu eksperymentalnego, którego efekty ocenić mamy. Kontrola tej skuteczności stanowi, standardowy krok w procedurze eksperymentalnej.

Niektóre eksperymenty społeczne nie dają oczekiwanego wyniku nie dlatego, że program eksperymentalny jest chybiony, lecz dlatego, że faktycznie nie został on wprowadzony w życie. Nowy styl zarządzania mógł być realizowany przez zru-

[91] Ch. A. Kiesler, Evaluating social change programs, w: G. Zaltman i in., Processes and phenomena of social changes, New York 1973..

tynizowanych kierowników w sposób niewiele różny od stylu tradycyjnego„ eksperymentalny program szkolny mógł być wprowadzany w życie przez niechętnych zmianie nauczycieli itd.

W metodologii eksperymentu zwraca się baczną uwagę na niezamierzone pozastymulacyjne wpływy eksperymentatorów na badanych (efekt- Hawthorne). Zjawisko to stawia problem rodzaju kontaktów między stronami biorącymi udział w eksperymencie. Wśród rozmaitych sposobów mających złu zaradzić ważne miejsce zajmuje ograniczenie tych kontaktów do minimum, chyba że wymaga tego sam badany problem. Rada ta odnosi się zarówno do eksperymentów laboratoryjnych, jak i do terenowych. Można ją odnieść z powodzeniem i do eksperymentów społecznych[92].

W eksperymentach prowadzonych na terenie jakiejś instytucji, takiej jak szkoła czy zakład przemysłowy, winien istnieć podział ról między personelem instytucji i badaczami pochodzącymi z zewnątrz. Cały eksperyment winien być zaplanowany przez obie strony. Jednak manipulacja powinna znajdować się w rękach personelu, odpowiednio do tego przeszkolonego. Badacze zaś zajmować się winni interpretacją i analizą wyników. Podział taki ma wiele dobrych stron. Ograniczony lub całkowicie wyeliminowany kontakt badaczy z badanymi uniemożliwia pierwszym pozastymulacyjne oddziaływanie na drugich a badanym uniemożliwia zorientowanie się, że zmiana której zostali poddani ma charakter eksperymentalny. Wszystko to ma nieocenione i dobroczynne następstwa dla „naturalności" działania instytucji i trafności uzyskanych wyników. Wspomniany podział ról sprawia i to jeszcze, że każda ze stron robi to, co umie najlepiej a zmiana eksperymentalna nie jest zakłócona przez równoczesną zmianę personelu.

Innym sposobem uniknięcia niepożądanego wpływu badaczy na instytucję jest oddziaływanie na nią w sposób pośredni: przez zmiany dokonywane przez wyższe szczeble organizacji, której instytucja jest elementem, przez zmiany w alokacji dóbr i przez działanie na środowisko, w jakim instytucja działa.

Kontrola zasięgu działania zmiany eksperymentalnej jest kolejnym zagadnieniem, które eksperymentatorzy muszą rozwiązać. Jednostki eksperymentalne i kontrolne tym się z definicji od siebie różnią, że jedne poddaje się bodźcom eksperymentalnym a drugie nie. Kontrola wymaga, by jedne i drugie były do siebie podobne a jednym ze sposobów tego ujednolicania jest dobór jednostek sąsiadujących ze sobą (przylegających do siebie obszarów, klas w jednej szkole, wydziałów w jednej fabryce, oddziałów w tej samej jednostce wojskowej itd.). Wtedy jednak może wystąpić zjawisko dyfuzji bodźca i efekt jego zatrze się. Problem ten stosuje się w całej rozciągłości i do eksperymentów społecznych[93]

Tu również zmiana musi mieć zasięg ograniczony. Może być to trudne z przynajmniej trzech powodów. Pierwszym jest komunikacja między jednostkami eksperymentalnymi i kontrolnymi. Powód drugi jest bardziej zasadniczy. Otóż program eksperymentalny uważany jest zwykle przez badanych za atrakcyjny. I słusznie, po cóż bowiem byłoby sprawdzać zmiany na gorsze. Jednostki przewidziane do grupy kontrolnej wywierają więc nacisk w kierunku rozciągnięcia programu od razu na szerszą skalę („my też chcemy być eksperymentalni!"), a to przeczyło by całej idei eksperymentu. Powód trzeci to kompensacyjne wyrównanie bodźca. Jeżeli program eksperymentalny polega m.in. na dostarczaniu pewnych szeroko rozumianych dóbr, to powstaje nierówność i na administratorów naciska się, by kompensowali grupom kontrolnym brak tych dóbr w jakiś inny niż eksperyment sposób. Jeśli ulegną, to innowacje eksperymentalne nie okażą swego wpływu, choćby mogły mieć znaczny. Wreszcie, w metodologii eksperymentowania zwrócono uwagę na istnienie wpływu uświadomionego braku bodźca na zachowanie się jedno-

[92] Zob. S. E. Seashore, Field experiments with formal organizations, w: W. M. Evan (ed.), Organizational experiments, New York 1971.

[93] Zob. Th. D. Cook, D. T. Campbell, The design and conduct of quasi-experiments and true experiments in field settings, w: M. D. Dunette (ed.), Handbook of industrial and organizational psychology, Chicago 1076.

stek kontrolnych[94]. Otóż jednostki te — zdarza się — bywają silnie motywowane do konkurencji z jednostkami eksperymentalnymi, a to w celu zredukowania spodziewanej różnicy między obiema grupami w wyniku działania bodźca. Jest tak np. wtedy, gdy pozytywny wynik eksperymentu godziłby w ich interesy lub poczucie wartości. Ale też może wystąpić i zjawisko odwrotne. Jeśli jednostki kontrolne uważają się za skrzywdzone w porównaniu z eksperymentalnymi, mogą — powodowane urazą — obniżyć jakość swego działania, np. wydajność pracy. I te ostatnie dwa względy musi mieć na uwadze ten, kto planuje eksperymentalne innowacje społeczne.

7.4. KONTROLA W EKSPERYMENTACH SPOŁECZNYCH

Kontrola służy — jak wiadomo — zapewnieniu eksperymentowi trafności wewnętrznej. Mówi się, że eksperyment jest wewnętrznie trafny, jeśli jego plan i sposób analizy wyników pozwala trafnie przypisać manipulacji wszystkie te i tylko te zmiany, które wywołała. Metodami kontroli są dobór odpowiedniego układu porównawczego dla układu eksperymentalnego albo — rzadziej — izolacja układu eksperymentalnego od zakłóceń, które utrudniłyby ocenę wpływu interwencji, gdyby się na nią nałożyły.

Kontrola — przypomnijmy — musi uwzględnić następujące czynniki: selekcję tj. istotne różnice między grupami zaangażowanymi w eksperyment; historię tj. wydarzenia w środowisku eksperymentu w trakcie jego trwania; dojrzewanie (czyli trend), na które bodziec może się tylko nałożyć; wpływ samego pretestu na badanych i posttest; zmianę narzędzi pomiarowych w trakcie eksperymentu, regresję statystyczną, która pojawia się wówczas, gdy grupy wybrane są ze względu na krańcowe wyniki pomiaru oraz wówczas, gdy bodziec wprowadza się w» chwili gdy wartości zmiennej zależnej znacznie odbiegają od normy; wreszcie „śmiertelność” tj. selektywne ubytki z grup zaangażowanych w eksperyment. Czynniki te zagrażają trafności eksperymentu i muszą być kontrolowane.

Wstępnym warunkiem trafnej oceny następstw sprawdzanych interwencji eksperymentalnych jest ujednolicenie grup eksperymentalnej i kontrolnej, czy też grup poddanych różnym wariantom manipulacji. Jeśli jednostek jest dużo a nadto nie są one zbyt zróżnicowane, pierwszorzędnym narzędziem ujednolicania jest losowy przydział jednostek do grup. Randomizacja gwarantuje brak systematycznych różnic między grupami i czyni mało możliwe wyjaśnienie różnic stwierdzonych po wprowadzeniu zmiany czy innym działaniem tej zmiany. Możliwości randomizacji istnieją w dziedzinach dzielących się na wiele względnie autonomicznych jednostek, stosunkowo przy tym do siebie podobnych. Wojsko, szkolnictwo, administracja niższego szczebla są przykładami takich dziedzin. Tam, gdzie jednostki są zróżnicowane, metodą ujednolica- > nia stać się powinno blokowanie i dzielenie losowe jednostek wewnątrz bloków na eksperymentalne i kontrolne. Ale w eksperymentach społecznych nieczęsto wykorzystuje się przewagi randomizacji. Zwykle dobiera się do układów eksperymentalnych układy, o których się sądzi, że są do nich najbardziej podobne ze względu na własności strukturalne, dynamiczne i relacje ze środowiskiem. Jest to dobre wyjście, jeśli układ eksperymentalny jest skomplikowany i niepowtarzalny. Lepsze są jednak takie układy porównawcze niż żadne. Ta preferencja nie zawsze jest podzielana przez eksperymentatorów społecznych.

Eksperymenty społeczne posługujące się układami porównawczymi, a zwłaszcza oparte na planach zrandomizowanych pozwalają radzić sobie z problemami kontroli. Grupy są wyrównane, podlegają tym samym zmianom zewnętrznym i trendom, zmiany wrażliwości narzędzi pomiarowych i wpływ pretestów jeśli ist-

[94] Ibidem.

nieją to we wszystkich grupach, jednakowe są też ubytki, regresja nie jest groźna, ponieważ dobór losowy ją wyklucza.

Co jednak robić, jeśli dla układu eksperymentalnego nie można znaleźć układu kontrolnego? Dwa sposoby postępowania można wówczas polecić.

Jeżeli program eksperymentalny jest tego rodzaju, że działa z różną mocą, w różnych okresach, to można uczynić porównania między tymi momentami. Grupa kontrolna i eksperymentalna będą wówczas złożone z odcinków czasu. Jest to mocny schemat, gdyż różnice między niedoskonale wyrównanymi jednostkami nie obciążają wyników, bowiem jednostki są te same. W ten chytry sposób Ross badał wpływ wprowadzonych w Anglii sankcji za prowadzenie samochodu w stanie nietrzeźwym na liczbę wypadków drogowych (nie było to badanie eksperymentalne, lecz studium reformy, ale jest to tutaj bez znaczenia, chodzi tylko o zilustrowanie logiki postępowania)[95] Porównywał on liczby wypadków w godzinach dojazdów do pracy z liczbami wypadków w noce weekendowe — prowadzenie samochodu w stanie nietrzeźwym za pierwszym razem zdarzało się rzadko, w drugim nagminnie. Okazało się, że silny spadek liczby wypadków wystąpił tylko w weekendowe noce, co potwierdzało hipotezę o skuteczności podjętej regulacji.

Jeżeli nie można wykorzystać nawet tego sposobu kontroli, to pozostaje jeszcze rozciągnięcie pomiarów na osi czasu, co stworzy ze schematu „pretest — posttest" schemat przerwanego szeregu czasowego o wiele odeń lepszy. Szereg czasowy jest zwykle dosyć łatwy do utworzenia, gdyż instytucje prowadzą okresowe oceny swej działalności i istnieją rozmaite statystyki. Można więc zrekonstruować rozwój jednostki eksperymentalnej w długim okresie poprzedzającym wprowadzenie interwencji eksperymentalnej. Łatwiej jeszcze będzie o dane obrazujące rozwój w długim czasie po interwencji. Istnieją specjalnie skonstruowane testy statystyczne pozwalające ocenić istotność przerwania ciągłości szeregu czasowego w momencie interwencji. Testy te nie są jednak w stanie dowieść, że czynnikiem, który zadecydował o nieciągłości, była interwencja eksperymentalna. Wydarzenia Współwystępujące z nią nie są więc w tym schemacie wykluczone jako przyczyny konkurencyjne. Bardzo często jednak szereg pozwala przekonać się, że zmiana jest bądź nie jest fragmentem długofalowego trendu, że jest bądź nie jest wynikiem regresji do średniej, że może być tylko zmianą pozorną, spowodowaną przez liberalizujące się lub zaostrzające kryteria klasyfikacji faktów, itd. Szereg czasowy pozwoli też stwierdzić, czy zmiana ma charakter przejściowy czy trwały.

Temu, kto chciałby wypróbować innowacje społeczne i nie mógłby skorzystać z układów kontrolnych, można udzielić rad następujących[96].

Nie należy wprowadzać programu eksperymentalnego w życie jakiegoś systemu społecznego w chwili jego kryzysu czy jego nadzwyczajnej koniunktury. Wartości zmiennych zależnych odbiegają wówczas od normy i prawdopodobnie nastąpi samorzutna ich regresja do średnich (w wyniku działania mechanizmów samostabilizujących). Lepiej poczekać do czasu unormowania się stanów systemu. Eksperymenty społeczne bywają odpowiedziami na nagły kryzys instytucji, tymczasem lepiej nadają się do szukania metod rozwiązywania problemów chronicznych.

Zmianę eksperymentalną należy wprowadzać totalnie i szybko. Inaczej trudno będzie odróżnić jej skutki od skutków innych zmian w środowisku i trendów długofalowych. Gdy wprowadzi się ją szybko, to ewentualne przerwanie ciągłości szeregu czasowego będzie wyraźne.

Należy unikać rozgłosu przed i w trakcie wprowadzania w życie programu eksperymentalnego. Nie należy np. poprzedzać go intensywną kampanią propagandową. Wprawdzie kampania może znakomicie podnieść skuteczność programu, ale o tyleż samo utrudni określenie tej skuteczności, a ono jest celem nadrzędnym eksperymentu. Stworzone przez kampanię zainteresowanie szerszych kręgów prowadzonym eksperymentem towarzyszy ponadto tylko nowościom, więc

[95] H. L. Ross, Law, science, and accidents: The British Road Safety Act of 1967, „Journal of Legal Studies" 1S73, vol. 2, nr 1.

[96] Zob.: M. W. Riecken, R. F. Boruch (eds.), Social experimentation. A method for planning and evaluating social intervention, New York 1974, s. 107 - 8.

stosowanie na szerszą skalę sprawdzonych w takich warunkach rozwiązań przebiegać będzie w sytuacjach o ten katalizator różnych, zatem możliwość uogólniania wyników eksperymentu zmniejszy się.

W czasie trwania eksperymentu nie należy zmieniać sposobu zbierania danych, zasad pomiaru i klasyfikacji faktów. Częstokroć reformy instytucji wiążą się z reformami i w tej dziedzinie, co daje wrażenie zmian, za którymi może nic się nie kryje. W eksperymentalnych próbach programów społecznych lepiej mieć wskaźniki gorsze, lecz stałe niż poprawiać ich jakość kosztem porównywalności zebranych danych.

Powyższe postulaty stosują się tak samo i do eksperymentów zaplanowanych na schematach z układami kontrolnymi, ale do eksperymentów bez takiej kontroli odnoszą się w stopniu szczególnym.

7.5. POMIAR NASTĘPSTW INNOWACJI EKSPERYMENTALNEJ

Dwie ważne sprawy w tym punkcie przedyskutować należy: rodzaj danych i pomiarów stosowanych do oceny efektywności programów społecznych oraz wybór momentu, w jakim pomiary te przeprowadzać należy.

Pomiar stanów rzeczy przed i po wprowadzeniu innowacji społecznych bardzo często nie wymaga nowych narzędzi, czynności i danych, gdyż systemy społeczne i tak prowadzą ciągłe lub okresowe pomiary swej struktury, funkcjonowania i efektywności. W szkole stawia się stopnie, w fabryce oblicza się wydajność pracy i płynność kadr, w prokuraturze prowadzi się statystyki przestępczości, itd. Dane zastane, pochodzące z instytucji, mają swe dobrze znane zalety: są już gotowe i są w sposób naturalny produkowane przez systemy. Ale równie dobrze znane są ich wady. Zasady ich zbierania często się zmieniają, przy czym zmiany te nie muszą być odpowiedzią na nowe problemy a mogą być dyktowane interesowną chęcią zapewnienia „sukcesu" eksperymentu. Dane urzędowe obciążone są często tendencyjnością, gdyż są podstawą oceny instytucji przez zwierzchników.

Kiedy badacz stawia sobie takie kryteria efektywności instytucji, których instytucja sama sobie nie stawia, to nie może on ograniczyć się do tego, co sama instytucja gromadzi w swych archiwach i bieżącej buchalterii. Zakłady pracy nie oceniają okresowo satysfakcji i frustracji, jakich doświadczają ich pracownicy. Pomiary, jakich dokonuje sam badacz, podzielić można na reaktywne i dyskretne[97]. Pierwsze dają badanym świadomość, że są przedmiotem badania i szczególnego zainteresowania, co jak wykazały już eksperymenty Eltona Mayo w zakładach Hawthorne może mieć wpływ na ich zachowanie, różny od wpływu samych manipulacji eksperymentalnych. Badania testowe i ankietowe należą do tej grupy, choć jeśli zbiorowość przywykła jest do nich, to reaktywność ich jest nieistotna. Z pomiarów dyskretnych wymienić trzeba przede wszystkim obserwację ukrytą.

Skuteczność eksperymentalnych innowacji społecznych bez wątpienia najwygodniej byłoby ocenić zaraz po ich wprowadzeniu. Nie trzeba byłoby rozciągać badań w czasie, szybko uzyskano by wyniki potrzebne do ulepszenia rzeczywistości, bliskie stany rzeczy pewniej można byłoby przypisać interwencji, zależności byłyby wyraźne. Jednakże efekty innowacji eksperymentalnej szczególnie istotne ze społecznego punktu widzenia trzeba mierzyć niejednokrotnie po długim czasie. Skuteczności resocjalizacji nie mierzy się poprawnością zachowania się wychowanków w domu poprawczym, lecz na wolności. Koncentracja na bezpośrednich następstwach zmiany eksperymentalnej prowadzi do przeceniania rozmiarów tych następstw, gdyż zwykle wiele z nich zanika w miarę upływu czasu. Stąd postulat opierania eksperymentów na schematach serii czasowych. Kryteria bezpośrednie

[97] Zob. E.J. Webb, D.T. Campbell, S. Sechresl, R.D. Schwartz, Unobtrusive measures: nonreactive research in the social sciences, Chicago I960.

są raczej wewnętrznymi kryteriami instytucji a odległe kryteriami, wedle których instytucje te oceniane są przez społeczeństwo.

Ocena sprawdzonej innowacji powinna — rzecz jasna — uwzględniać jej skutki pozytywne i negatywne, zamierzone i niezamierzone; możliwe są ich rozmaite kombinacje.

7.6. TRAFNOŚĆ ZEWNĘTRZNA EKSPERYMENTÓW SPOŁECZNYCH

Drugim — obok trafności wewnętrznej — podstawowym kryterium oceny eksperymentu jest trafność zewnętrzna. Nie wystarczy z niezbitą pewnością ustalić skutki manipulacji. Trzeba nadto znać zakres ogólności ustalonych związków i prowadzić eksperymenty tak, by zakres ten był szeroki. Od czego więc zależy możliwość uogólniania wyników eksperymentów społecznych czyli możliwość stosowania sprawdzonych rozwiązań na szeroką skalę?

1. Wyniki eksperymentu społecznego są tym bardziej ogólne, im bardziej jednostki eksperymentalne są reprezentatywne dla szerszych układów społecznych. Wiadomo, że nie zawsze takimi bywają, przy czym czynniki ze względu na które jednostki są niereprezentatywne, działają na korzyść bodźca, wyjątkowo tylko jest inaczej. Przyczyną tego może być samoselekcja: na poddanie się eksperymentowi godzą się te instytucje, które są bardziej dynamiczne, chłonne na innowacje itd. Może nią być też tendencyjny dobór: na jednostki eksperymentalne wybiera się takie, o których sądzi się, że eksperyment się w nich „uda". Radą na to wszystko jest dobór losowy, zapewniający reprezentatywność jednostek eksperymentalnych, dobór jednostek eksperymentalnych spośród jednostek modalnych w zbiorowości i celowe różnicowanie grup eksperymentalnych tak, by móc sprawdzić, czy stwierdzone zależności załamują się na tych zróżnicowaniach.

2. Możliwość uogólniania wyników tego eksperymentu jest tym większa, im bardziej sytuacje eksperymentalne były typowe dla sytuacji pozaeksperymentalnych. Dla oceny tego wymiaru istotne są sprawy następujące: Jednostka eksperymentalna działa w środowisku jednostek tradycyjnych, co może w szczególny sposób ułatwić lub utrudnić jej funkcjonowanie i tym samym ograniczyć możliwości uogólniania. Na trudności te zwracają uwagę ekonomiści piszący o eksperymentalnych próbach zmian gospodarczych. „Wynik takiego «eksperymentu» nie może być niestety wystarczającą podstawą do zaniechania lub upowszechnienia danej innowacji — pisze Beksiak. Jeśli nie była ona sprzeczna z «resztą» istniejącego systemu, to z tego wcale nie wynika, że zastosowana powszechnie nie przyniesie ujemnych skutków. Jeśli zaś innowacja ta nie została zaadaptowana obecnie, nie dyskwalifikuje jej to całkowicie. Być może dopiero całościowa zmiana systemu kierowania gospodarką pozwoliłaby na zastosowanie tego właśnie narzędzia"[98].

Wyniki eksperymentów społecznych często obciążone są efektem Hawthorne. Wprowadzaniu eksperymentalnych innowacji społecznych towarzyszą szczególne stany emocjonalne, współdecydujące o skuteczności wypróbowywanych rozwiązań. Przy szerokim wprowadzeniu w życie tych praktyk, efekt ten by znikł.

W wielu eksperymentach społecznych personel zaangażowany w realizację nowego programu jest szczególnej jakości. Przedsięwzięcia te są często prowadzone przez kadrę wyżej kwalifikowaną i silniej motywowaną niż kadra, która będzie wprowadzać sprawdzone innowacje w życie. Jeśli od jakości kadry zależy skuteczność programu, to ogólność wniosków wyciągniętych z takich eksperymentów się kurczy.

3. Ogólność wyników eksperymentów społecznych zależy od tego, na ile powszechnie kryteria stosowane przez badaczy do oceny programu eksperymental-

[98] J. Beksiak, Społeczeństwo gospodarujące, Warszawa 1976, s. 142.

nego stosowane są przez społeczeństwo do oceny rozwiązań wchodzących w grę problemów społecznych. Ten punkt nie wymaga objaśnienia.

7.7. PRZYKŁAD —EKSPERYMENT RESOCJALIZACYJNY

Wiele problemów podniesionych w dotychczasowej dyskusji zilustrujemy na przykładzie interesującego eksperymentu w zakresie resocjalizacji młodocianych przestępców, prowadzonego pod koniec lat pięćdziesiątych w Szczypiornie przez Katedrę Pedagogiki Ogólnej Uniwersytetu Warszawskiego i Katedrę Prawa Karnego Uniwersytetu Wrocławskiego[99]

Decyzję o podjęciu tego eksperymentu poprzedziło zdanie sobie przez badaczy sprawy z miernej skuteczności dotychczasowych metod resocjalizacji oraz z istnienia wielu sprzecznych poglądów na temat tych metod, przy równoczesnym braku koncepcji sprawdzonych empirycznie.

Istniejące praktyki reedukacji młodocianych przestępców w zakładach zamkniętych polegały na zaprowadzeniu w tych zakładach rygoru więziennego. Rygor ten miał charakter zewnętrzny i przeciwstawiał więźniów personelowi. Miał on też charakter represyjny i zamiast resocjalizować, wypaczał osobowości młodocianych. Autorzy eksperymentu wyszli natomiast z założeń przeciwnych. Uznali, że należy „dotrzeć do osobowości więźnia, by resocjalizacja odbywała się z jego udziałem" a „system więzienny winien wprowadzać w życie społeczne a nie wypaczać osobowości".

Konkretyzacja tych założeń ogólnych doprowadziła autorów do projektu takiego zakładu zamkniętego, którego wygląd nie będzie przypominał więzienia, w którym zniesie się zbędne ograniczenia„ w którym stosunek administracji do wychowanków będzie ojcowski i wytworzy się uczuciowy związek wychowanków z zakładem, w którym rozwinie się życie zespołowe oparte na grupach koleżeńskich, zindywidualizuje się metody wychowawcze, będzie się liczyło z zainteresowaniami i poziomem umysłowym wychowanków, stawiać się im będzie coraz wyższe wymagania, ale i przyznawać coraz większe prawa, wyuczy się ich zawodu, umożliwi kontakt z rodziną i szerszymi sprawami społecznymi, oswoi z wolnością i będzie się utrzymywać kontakt z wychowankami po ich wyjściu na wolność.

Hipoteza eksperymentalna brzmiała: „O ile zostaną spełnione te wymogi, to społeczne życie zakładu stanie się właściwym środowiskiem, w którym dokonywać się będzie żmudny i trudny proces resocjalizacji przestępcy". Mowa tu o resocjalizacji skuteczniejszej niż w zakładach urządzonych tradycyjnie. Kryteriami powodzenia resocjalizacji były: rozwój życia społecznego w zakładzie, wzrost dyscypliny wewnętrznej wychowanków, podejmowanie trudnych, samodzielnych decyzji a nade wszystko społecznie aprobowany charakter zachowań po wyjściu na wolność, w szczególności stopa recydywy.

W celu sprawdzenia tej hipotezy utworzono specjalny zakład wychowawczy w Szczypiornie, będący realizacją projektu eksperymentalnego. W zakładaniu tej instytucji zwracają uwagę dwa momenty: dobór wychowanków i dobór personelu. Wybrano więźniów szczególnie trudnych, o małych szansach resocjalizacji, ale zarazem będących pewnymi indywidualnościami. Zgodnie z regułami dedukcyjnej krytyki hipotez fakt ten czynił z eksperymentu test surowy, szczególnie cenny dla hipotezy — gdyby nowy system sprawdził się w tej grupie, należałoby oczekiwać, że tym bardziej sprawdzi się w grupach przestępców mniej zdemoralizowanych. Szczególnej jakości był też personel — „osobowość i praca tych ludzi nadała indywidualne piętno zakładowi w Szczypiornie i faktycznie zadecydowała o dodatnich wynikach eksperymentu" — tu dla odmiany specyfika personelu, osłabia wybitnie

[99] I-I. Świda, W. Świda, Młodociani przestępcy w więzieniu, Warszawa 1961, stąd cytaty na s. 180 - 181, 182.

stosowalność sprawdzonych rozwiązań w warunkach typowych, choć i wskazuje jasno czego wymaga nowe urządzenie.

Dla systemu eksperymentalnego nie dobierano określonego systemu porównawczego. Zakładano bowiem, słusznie zresztą, że jest on zbędny, gdyż ani w środowisku zewnętrznym, ani w osobowościach wychowanków nie wystąpią nagle czynniki, które mogłyby dać , powód do wyjaśnień konkurencyjnych wyników eksperymentu. Nie dopuszczono też istnienia ogólniejszego trendu podnoszenia się skuteczności działań resocjalizacyjnych, którym można ’ byłoby tłumaczyć zaobserwowane zmiany w sposób alternatywny. O samych objętych eksperymentem wychowankach zakładano natomiast milcząco, że gdyby pozostali oni w poprzednich zakładach, to właśnie zgodnie z powszechnymi tendencjami nie poczyniliby większych postępów na drodze resocjalizacji.

Zaobserwowano interesujące zjawisko wysuwania przez wychowanków roszczeń z tytułu eksperymentu — mieli oni świadomość, że ich sytuacja jest eksperymentalna, zatem wyjątkowa. Okoliczność ta jest kontaminacją względem resocjalizującego działania nowych warunków i dla dobra trafności wyników eksperymentu oraz ogólności dyrektyw socjotechnicznych nie powinna była mieć miejsca.

Eksperyment planowany był na wiele lat, lecz został raptownie przerwany po ucieczce grupy więźniów z wysokimi wyrokami — stosunkowo łatwiej w warunkach dużego zaufania a małego rygoru ze strony administracji zakładu. Nadrzędne organa administracji więziennej uznały ucieczkę za dowód niepowodzenia eksperymentu.

Zdaniem autorów (zgadzamy się z nimi w zupełności) ewaluacja wyników eksperymentu społecznego takiego jak opisywany jest zagadnieniem bardziej złożonym. Wskazują oni na fakt, że brak prób ucieczek nie jest jedyną miarą skuteczności systemu wychowawczego więzienia. Zaobserwowali oni wystąpienie innych przewidywanych przez testowaną hipotezę pozytywnych objawów. Rozwinęło się życie społeczne i samorządowe, u większości wychowanków przejawiało się poczucie rzeczywistości i samodyscyplina. Ostateczny sprawdzian skuteczności oddziaływania resocjalizującego mógłby nastąpić dopiero po dłuższej obserwacji zachowania się w zakładzie a nade wszystko po wyjściu na wolność, na co eksperyment trwał zbyt krótko. Niemniej jednak autorzy przewidują na podstawie obserwacji, których dokonać zdążyli, że „dalsza praca wychowawcza mogłaby uczynić z wielu chłopców prawdziwie wartościowych ludzi — odzyskanych dla społeczeństwa”. Uważają oni generalnie, że eksperyment dowiódł, że istnieje możliwość reedukacji młodocianych przestępców, nawet bardzo zdemoralizowanych a obrany kierunek działania resocjalizacyjnego jest Kierunkiem zasadniczo słusznym.

Przypadek ten ujawnia pewne kłopoty ewaluacji eksperymentów społecznych. Pierwszy bierze się z istmienia wielu niewspółmiernych kryteriów skuteczności sprawdzanych rozwiązań. Drugi pochodzi z długiej skali czasu, jakiej wymaga się do oceny jego wyników. W tej sytuacji trudno o oceny jednomyślne i jednoznaczne. Bezpośrednie skutki działania nowych urządzeń mogą być pozytywne ze względu na jedne kryteria, a negatywne ze względu na inne. Dalekosiężne, jeśli nie znikną w ogóle, mogą okazać się odwrotne.

Wartość eksperymentów społecznych mierzy się powszechnością skutecznych zastosowań ich wyników. Jeśli tych się nie stosuje, to ich wartość praktyczna jest z definicji żadna; chyba że bogacą się przy ich okazji teorie społeczne, ale to jest korzyść trudna do osiągnięcia. Co się tyczy eksperymentu w Szczypiornie, to — o ile nam wiadomo — jego wyniki poszły w niepamięć a system resocjalizacji młodocianych przestępców w zakładach zamkniętych budzi nadal bardzo poważną krytykę.

7.8. KORZYŚCI I TRUDNOŚCI EKSPERYMENTÓW SPOŁECZNYCH

Korzyści z eksperymentalnych prób zmian społecznych czerpać może zarówno praktyka, jak i nauka teoretyczna. Korzyści praktyczne są jednak pierwsze, korzyści dla nauki są niejako produktem ubocznym eksperymentowania społecznego.

Eksperyment społeczny sprzyja podejmowaniu trafnych decyzji praktycznych, wywołaniu pożądanych zachowań i zmiany społecznej. Informując o prawdopodobnych następstwach alternatyw stojących przed decydentem, zmniejsza ryzyko kosztownych nieraz pomyłek. Zmniejsza społeczne koszty chybionych decyzji, małe w skali systemu eksperymentalnego, duże w skali makrospołecznej.

Eksperyment społeczny jest dobrą strategią wprowadzania zmiany społecznej. Posiada ona takie cechy, które pomagają innowacjom rozprzestrzeniać się. Innowacje przyjmowane są tym łatwiej, im łatwiejsza do wykazania jest wartość nowej praktyki, im łatwiej można przywrócić stan rzeczy, jaki istniał przed przyjęciem innowacji, im mniejszych zasobów innowacja wymaga; nie są to oczywiście wszystkie własności innowacji sprzyjające ich przyjmowaniu. Eksperyment posiada wymienione własności: uwidacznia zalety innowacji, jego skutki są odwracalne, wymaga ograniczonych zasobów.

Eksperymenty społeczne pozwalają w szczególny sposób połączyć dwa ekstrema działania praktycznego: śmiałość i radykalizm z ostrożnością i powściągliwością. Dzięki eksperymentalnym próbom zmian społecznych radykalizm nie musi przeradzać w się w kosztowne awanturnictwo, a powściągliwość nie musi wyradzać się w jałowy konserwatyzm. Test eksperymentalny pozwala na dokonywanie w małej skali nawet daleko idących zmian społecznych. Konserwatysta dbały o trwałość instytucji przekona się, czy zmiany te istotnie jej zagrażają i czy wynik ich jest istotnie niepewny. Radykał dowie się, czy jego program da się zrealizować i czy się opłaci. Sytuację zanotowaną w tym punkcie można zaobserwować często: jeśli istnieje opór przed innowacją, to wiele bardzo nawet śmiałych nowinek przemycić można pod nagłówkiem „eksperyment", a zarazem zmniejszyć wydatnie koszty ewentualnych pomyłek.

Wyniki eksperymentów społecznych posiadają dużą siłę perswazyjną i wartość rozstrzygającą względem sprzecznych przekonań społecznych. W socjologii przekonania takie są rzeczą nagminną, wobec czego praktycy mogą przebierać stosownie do potrzeb i upodobań w hipotezach, domysłach i uzasadnieniach różnych sposobów działania. Przebieranie to można zastąpić praktycznymi testami konkurujących ze sobą rozwiązań. Siła wniosków eksperymentalnych nie musi jednak zachęcać do eksperymentowania. Eksperyment jest bezlitosnym sposobem oceny wartości rozwiązań praktycznych — ci administratorzy, którzy

obawiają się o wartość swych praktyk, wybiorą takie metody, których wyniki będą mogli wielorako interpretować.

Podejmowanie decyzji praktycznych o przyjęciu tej a nie innej metody działania dopiero po uprzednim eksperymentalnym jej wypróbowaniu upodabnia podejmowanie decyzji praktycznych do procesu kontroli hipotez naukowych. Proces kontroli hipotez naukowych ma charakter intersubiektywny: każdy badacz posiadający odpowiednie kompetencje może go kontrolować przez powtórzenie, nadzór czy nade wszystko przez samą możliwość powtórzenia. Dzięki temu nauka ma charakter publiczny. Polityka społeczna oparta na sprawdzianach eksperymentalnych staje się tak jak nauka dziedziną jawną, poddaną _ w pewnej mierze publicznej kontroli.

Eksperymenty społeczne mogą dostarczyć materiałów dla teoretycznej analizy zjawisk społecznych, takich materiałów, które w inny sposób znaleźć niełatwo. Jeśli badacz potrafi zinterpretować w terminach ogólnych wprowadzane innowacje, ich następstwa i czynniki sytuacyjne, od których zależy ich pojawienie się, to eksperyment społeczny stanie się nie tylko narzędziem zmiany społecznej, ale i budowy teorii socjologicznej. Np. eksperymentalne badania efektywności nowych

form zarządzania mogłyby przyczynić się do rozwoju teorii organizacji. Musimy jednak liczyć się z trudnościami takich interpretacji. Manipulacji w eksperymentach społecznych nie dobiera się ze względu na ich doniosłość teoretyczną, lecz wagę praktyczną — z teoretycznego punktu widzenia mogą one po prostu być mało ciekawe. Manipulacje te mają charakter tak złożony, że trudno wskazać, jakie ich aspekty wywołują zmianę. Wreszcie, w eksperymentach społecznych ustala się pewne związki, relatywizując je przy tym do bliżej nie sprecyzowanego a tylko wymienionego z nazwy kontekstu — po to, by związki te zinterpretować teoretycznie, trzeba byłoby najpierw zinterpretować w ten sposób owe warunki.

Trzeba jednak powiedzieć również o trudnościach eksperymentowania społecznego. Zaznaczymy je tylko, gdyż o wielu problemach, jakie wyrastają przed eksperymentatorami, już mówiliśmy.

Eksperymentowanie *in vivo* wymaga kontrolowania pewnych odcinków życia zbiorowego, toteż badania tego typu mogą być tylko wspólnym dziełem administratorów i badaczy. Ich porozumienie się może być sprawą trudną. Administratorzy musieliby bowiem zgodzić się na zmianę zasad funkcjonowania podległych im instytucji zgodną z wymaganiami schematu eksperymentalnego. Tymczasem eksperymenty społeczne są pewnym rodzajem zmiany, napotykają więc ten rodzaj oporów, na jaki natrafiają wszelkie innowacje godzące w istniejący układ statusów, interesów i nawyków. W dodatku, selektywne wprowadzanie innowacji, ograniczenie jej zasięgu do pewnych tylko, a zwłaszcza tylko do wylosowanych elementów systemów może być trudne do usprawiedliwienia. Jeśli bowiem innowacja jest dobra, to należy wprowadzić ją w całym systemie a jeśli jest zła, to nie należy jej wprowadzać w ogóle — może argumentować administrator.

Z wymienionych względów łatwiej eksperymentować w sytuacjach, gdy planowana innowacja nie może być wprowadzana w całym systemie od razu, chociażby z powodu braku środków, instruktorów, możliwości jej nadzoru itd. Wówczas można łatwiej porozumieć się co do konieczności takiej alokacji innowacji, by jednostki nią objęte i jednostki pozostałe były wyrównane pod istotnymi względami. Wówczas nie tylko zyska się możliwość nadzoru nad wprowadzaniem innowacji, lecz również sprawdzi się ją i wyniki tego sprawdzenia będzie można zastosować w kolejnych etapach procesu wprowadzania innowacji.

Łatwiej również eksperymentować w tzw. instytucjach totalnych — takich, które obejmują kontrolą całość zachowań i kontaktów swych członków; zakład resocjalizacyjny jest przykładem takiej instytucji. W instytucjach totalnych postulaty kontroli eksperymentalnej realizowane są niezależnie od potrzeb eksperymentatorów; jest to kontrola przez izolację.

Łatwiej eksperymentować także wówczas, gdy tworzy się jakieś instytucje od podstaw albo przeprowadza się gruntowne reformy instytucji istniejących. W takich sytuacjach zmiana i tak się dokonuje a uwzględnienie potrzeb eksperymentatora może wymagać niewielkich dodatkowych przedsięwzięć.

Wymieńmy i inne, niż brak władzy, trudności eksperymentowania społecznego.

Integralne systemy społeczne stawiają opór zmianom cząstkowym, a eksperymenty społeczne są takimi właśnie zmianami. Eksperymenty społeczne są badaniami kosztownymi, na co jednak łatwo można odpowiedzieć, że metoda prób i błędów jest o wiele bardziej kosztowna. Ewaluacja programów eksperymentalnych wymaga długiego czasu, często tak długiego, że potrzeby i warunki środowiska zmieniają się istotnie. Powinno się więc planować i wykonywać eksperymenty społeczne wystarczająco wcześniej, a nie wtedy, gdy potrzeba rozwiązania problemu społecznego stanie się palącą. Długotrwałość eksperymentu społecznego sprawia zresztą, że jedną z metod torpedowania naglącej zmiany jest skierowanie jej do eksperymentalnego sprawdzania.

Zwróćmy wreszcie uwagę na dosyć specjalną trudność prowadzenia eksperymentów społecznych. Jest nią tradycyjny sposób pojmowania, roli badacza jako rejestratora, analizatora i interpretatora faktów niezależnie odeń istniejących. Podejmowanie eksperymentalnych programów społecznych wymagałoby istotnej

reorientacji tej roli w stronę działania. Wymagałoby również ściślejszej niż dotąd współpracy przedstawicieli różnych dyscyplin społecznych.

*

* *

Kończąc rozważania poświęcone eksperymentowaniu społecznemu, zanotujmy śmiały pogląd upatrujący w prowadzeniu i wykorzystaniu eksperymentów społecznych szansę przezwyciężania przez nauki społeczne zubożającego je skrępowania materiałami tylko zastanymi i obserwacyjnymi. Próbując odgadnąć metodologiczną przyszłość nauk społecznych, Stanisław Ossowski pisał: „A może socjolog empiryk, ograniczający swą dyscyplinę wyłącznie do badań nad bezpośrednio dostępną mu rzeczywistością, otworzy sobie drogę ku szerszej problematyce i szerszym uogólnieniom, kompensując brak materiałów historycznych przez stosowanie wielorakich eksperymentów społecznych rozległej skali? Pomysły takich eksperymentów mają znacznie starszą tradycję, niż się zwykle mniema. Jeżeli wierzyć Herodotowi, o możliwości eksperymentów społecznych dla celów naukowych wiedział już ów faraon, który na drodze eksperymentalnej chciał zaspokoić swe zainteresowania lingwistyczno-socjologiczne, dotyczące naturalnego języka. Tylko dzisiaj pole dla pomysłów jest nieporównanie szersze niż za czasów owego faraona i nieporównanie szersze są możliwości ich realizacji”[100]

[100] Tak kończy się dzieło Ossowskiego O osobliwościach nauk społecznych.

CZĘŚĆ III GRANICE EKSPERYMENTU W BADANIACH SPOŁECZNYCH

Jedną z bardziej rzucających się w oczy osobliwości badań społecznych są niewielkie w stosunku do badań przyrodniczych możliwości eksperymentowania. Widocznie nauki społeczne napotykają w realizacji swych zadań poznawczych pewne przeszkody dla eksperymentu, nieznane przyrodoznawstwu w ogóle' w takiej postaci czy natężeniu. Celem tej części pracy jest właśnie opisanie owych barier. Zadanie to jest równoznaczne z udzieleniem odpowiedzi na pytanie: Dlaczego w badaniach społecznych stosuje się eksperyment w zakresie tak wąskim, mimo że jest on metodą tak doskonałą?

Aby zadanie to sobie ułatwić, wprowadzę pojęcie Idealnego Eksperymentatora. Z definicji jest ono fikcją metodologiczną, ale dlatego właśnie jest użyteczne, gdyż pozwala przez opozycję zdać sobie sprawę z rozziewu, a nawet przepaści, między możliwościami badacza a możliwościami metody. Idealny Eksperymentator obdarzony byłby ludzką co prawda ciekawością, ale nadludzką władzą i nieludzką moralnością. Mógłby on więc badać eksperymentalnie wszystkie zjawiska i sprawdzać eksperymentalnie wszystkie hipotezy. Mógłby wpływać na naturalny , tok rozwoju systemów społecznych i osobowości: cofać go i przyśpieszać w tempie dowolnym, odmieniać jego kierunki, zmieniać zasady funkcjonowania. Mógłby reprodukować zjawiska społeczne w kombinacjach i ilościach dowolnych, wypreparować z rzeczywistości społecznej rozmaite zjawiska i umieszczać w innych warunkach, izolować jedne zjawiska od drugich, wzmagać działanie jednych czynników a osłabiać innych itd. Żadne granice eksperymentowania by dla niego nie istniały. Istnieją one natomiast dla eksperymentatorów rzeczywistych.

8. TECHNICZNE GRANICE EKSPERYMENTU

„Eksperymentować można […] tylko w zakresie tych zjawisk, które potrafimy bądź sami wywołać, bądź sami dowolnie je zmieniać" — stwierdza krótko Ajdukiewicz[101]. W rozdziale niniejszym omówimy ograniczenia eksperymentowania wynikające właśnie z braku możliwości wywoływania lub dowolnego zmieniania zjawisk. Nazwiemy je granicami technicznymi.

8.1. HISTORYCZNY CHARAKTER ZJAWISK SPOŁECZNYCH

Wśród zjawisk będących przedmiotem badania w naukach społecznych pokaźną choć niewymierną część stanowią zjawiska historyczne. Określenie to obejmuje zjawiska minione, należące do przeszłości a także zjawiska co prawda współczesne, lecz w dużym stopniu wyznaczone przez pierwsze. Zakres sytuacji problemowych, w których konieczne jest badanie zjawisk obu tych rodzajów jest pierwszym progiem ograniczającym możliwości eksperymentu w badaniach społecznych.

Zjawiska minione występują w naukach społecznych w paru rolach. Są one samodzielnym przedmiotem badań — tak zajmuje się, nimi historia rozumiana po-

[101] K. Ajdukiewicz, Logika pragmatyczna, Warszawa 1975, s. 229.

tocznie: nauka ta usiłuje opisać i wyjaśnić zdarzenia przeszłe, gdyż te same w sobie wydają się jej godne poznania. Są też koniecznym budulcem dla teorii społecznych tak teorii średniego zasięgu, jak też teorii rozwoju społecznego. Wreszcie, fakty przeszłe występują w roli wyjaśniającej wobec faktów współczesnych — ich zbadanie jest nieodzowne dla ustalenia genealogii teraźniejszości. Zawsze jednak o dyskutowanych zjawiskach najlepiej powiedzieć, że ich już nie ma, zatem bezpośredni z nimi kontakt, warunek *sine qua non* eksperymentu, jest wykluczony — można je tylko inferować ze świadectw.

Tymczasem eksperyment zastosowany w historii byłby narzędziem wybornym, pozwalałby rozstrzygnąć liczne kwestie nierozwiązane, a zwłaszcza rozwiązane rozmaicie. W szczególności znalazłyby się wśród nich absorbujące umysły historyków pytania o historyczną rolę wielkich wydarzeń, wybitnych jednostek i idei dziejotwórczych oraz o przyczyny innych wydarzeń, uprzednio podane hipotetycznie. Intensywność sporów wokół nich toczonych jest — z metodologicznego punktu widzenia — następstwem trudności znalezienia w historii rozumianej jako proces sytuacji spełniających kanony Milla, różnych pod jednym tylko względem — istnieniem bądź brakiem wydarzenia, procesu czy idei, których skutki chce się określić. Fikcyjny Idealny Eksperymentator nie miałby tego typu kłopotów i łatwo posprawdzałby wszystkie konkurencyjne hipotezy i domysły.' Chcąc precyzyjnie określić historyczną rolę jakiejś indywidualności, wynalazku czy procesu powtórzyłby on historię raz jeszcze z jednym jedynym tylko wyjątkiem, że postać ta nie urodziłaby się w ogóle, nie dokonano by wynalazku, nie nastąpiłby proces. Okazało by się wtedy, czy ich role odegrałyby z powodzeniem inne wydarzenia i inni aktorzy historii. Chcąc zaś określić wagę jakichś elementów podłoża historycznego w rozwoju społeczeństwa lub instytucji, Idealny Eksperymentator stworzyłoby inne społeczeństwo i instytucję o podłożu podobnym pod wszystkimi istotnymi — poza badanym — względami.

Nie istnieją jakiekolwiek możliwości powtarzania biegu dziejów, toteż historia jest dyscypliną *par excellence* nieeksperymentalną. Historykowi pozostaje wyszukiwanie w niezależnej odeń rzeczywistości sytuacji możliwie bliskich wymaganym przez schematy indukcji — na tym m.in. polega tzw. metoda historyczno-porównawcza.

Zanotujemy jednak dla ścisłości, że nawet i w naukach historycznych istnieją procedery nazywane eksperymentami.

Często historyk modyfikuje proces historyczny we własnej wyobraźni. Chcąc określić następstwa pewnego wydarzenia, eliminuje on w myśli owo wydarzenie z realnego procesu historycznego i próbuje opisać hipotetyczny przebieg procesu przy założeniu niezajścia tego wydarzenia. Zabiegi takie nazywa się eksperymentem myślowym. Dokonuje się on w umyśle badacza. Dzięki temu, że tylko w umyśle, jest on — przestrzega Weber — „zabiegiem niebezpiecznym i niepewnym"[102] — dzięki zaś temu, że umysł, w którym on się odbywa, należy do badacza, nie jest on zabiegiem lak arbitralnym i jałowym, jak mogłoby się wydawać.

Klasycznym i wysłużonym przykładem eksperymentu myślowego w naukach historycznych jest piękna analiza znaczenia zwycięstwa pod Maratonem dla kultury europejskiej; można ją znaleźć w dziełach Webera i Nagla[103]. Polegała ona na szukaniu empirycznie uzasadnionej odpowiedzi na pytanie, jak wyglądałaby historia naszej kultury, gdyby zwycięstwo wzięli wówczas Persowie. Ale zabiegi tego rodzaju obecne są nie tylko w traktatach metodologów — są one zgoła pospolite w pracach historyków.

Postawmy pytanie o wartość i przydatność takich zabiegów. W jaki sposób eksperyment myślowy może — bez obserwacji — prowadzić do lepszego poznania rzeczywistości?

[102] M. Weber, The theory of social and economic organization, New York 1947, s. 97.

[103] Zob. M. Weber, The methodology of the social sciences, Glencoe 1949, s. 174-5; E. Nagel, Struktura nauki, Warszawa 1970, s. 504 - 6.

Eksperymenty myślowe oparte są na formule kontr- faktycznego okresu warunkowego. Są one tyle warte, ile warte są prawa empiryczne, przy pomocy których z poprzednika takiego okresu wywodzi się jego następnik.

W przeciwieństwie do eksperymentu rzeczywistego, w którym manipuluje się zjawiskami, obserwuje się skutki manipulacji i stwierdza się nowe zależności, eksperyment myślowy polega na manipulowaniu wyłącznie posiadaną wiedzą i wywodzeniu przy pomocy istniejących zależności następstw tych manipulacji. Nie może on więc pełnić podstawowej roli eksperymentu tj. kontrolować hipotez przez ich konfrontację z faktami. Rola jego polega jednak na czym innym: pozwala on lepiej uświadomić sobie znaczenie wydarzenia kontrfaktycznie zaprzeczanego. Określenie hipotetycznych skutków niezajścia pewnego wydarzenia i konfrontacja hipotetycznego procesu historycznego z procesem rzeczywistym ułatwia zdanie sobie sprawy z roli tego wydarzenia w realnym procesie historycznym.

Możemy więc powiedzieć, że eksperyment myślowy wcale eksperymentem nie jest — jest on tylko pewnym rodzajem rozumowania.

Niekiedy historycy podejmują próby „powtórzenia" pewnych wydarzeń historycznych. Najbardziej znanym z takich eksperymentów jest niewątpliwie wyprawa „Kon-Tiki"[104]. Jej celem było wykazanie możliwości zasiedlenia Polinezji przez dawnych mieszkańców Ameryki. Wyprawa dowiodła, że środki komunikacyjne, jakimi oni dysponowali były do tego wystarczające. Nie dowiodła ona i dowieść nie mogła niczego więcej ponad tę możliwość. Ale przeciwnicy teorii o amerykańskim pochodzeniu Polinezyjczyków powoływali się na niemożliwość pokonania wielkich przestrzeni oceanicznych na prymitywnych tratwach. Sukces wyprawy „Kon-Tiki" argument ten wytrącił im z ręki. Istnieje cały kierunek badań archeologicznych zwany archeologią doświadczalną — jego entuzjaści wykonują bardzo nieraz ciekawe eksperymenty sprawdzające działanie dawnych technik uprawy gleby, transportu, obróbki surowców, wytwarzania narzędzi itd.[105] Eksperymenty te wnoszą niejednokrotnie wiele do znajomości przeszłości. Czy jednak są to eksperymenty z zakresu społecznych nauk historycznych? Wątpliwe, są one eksperymentami z zakresu agrotechniki, sztuki żeglowania, metalurgii czy budownictwa, tyle że techniki w nich badane należą już do przeszłości.

Tak więc, z dwu procedur badawczych stosowanych w naukach historycznych, a nazywanych eksperymentami, pierwsza nie jest w ogóle eksperymentem a druga nie jest eksperymentem z zakresu społecznych nauk historycznych. Wehikuł czasu nie istnieje.

Wśród uwarunkowań zjawisk społecznych wyróżnia się przyczyny bliższe'(bezpośrednie) i dalsze. Jakkolwiek nieprecyzyjna i dowolna byłaby ta dystynkcja, to wiążą się z nią dwa typy wyjaśnień: przyczynowe i funkcjonalne — odwołujące się do bezpośrednich przyczyn i skutków oraz genetyczne — szukające przyczyn dalszych, odtwarzające stadia formowania się zjawiska. Istnieje zgoda co do tego, że są to podejścia komplementarne: struktury funkcjonalne i konfiguracje przyczynowe same są wytworem historii i biografii. Niemniej jednak preferencja dla jednego bądź drugiego rodzaju wyjaśnień jest jednym z kryteriów zróżnicowania wewnątrz nauki. Tak jest w antropologii, gdzie istnieje podejście funkcjonalistyczne i ewolucjonistyczne. Tak jest i w psychologii, w której wyróżnia się interpretacje rozwojowe i sytuacyjne.

Każde z tych podejść inaczej widzi rolę eksperymentu. Jeśli zachowania rozumie się jako reakcje na bodźce sytuacyjne, to po to, aby sprawdzić ich związek, trzeba zmierzyć zachowania przed i po działaniu bodźców — ale to jest przecież opis metody eksperymentalnej! Staje się ona podstawową metodą badań sytuacyjnych. Jeśli natomiast wyznaczniki zachowania widzi się w przeszłości osobników i systemów społecznych, to chcąc ustalić owe wyznaczniki, sięgnąć trzeba raczej po inne narzędzie, choćby wywiad docierający do przeżyć z okresu dzieciństwa czy

[104] Zob. T. Heyerdahl, Wyprawa Kon-Tiki, Warszawa 1972.
[105] Zob. J. Coles, Archeologia doświadczalna, Warszawa 1977.

analizy historyczne. Tak więc preferowanie historycznych wyjaśnień zachowania ogranicza zarazem przydatność eksperymentu.

Ta bariera działa również w naukach przyrodniczych. Historia w znaczeniu popularnym nie ma bowiem monopolu na badanie przeszłości. W rodzinie nauk przyrodniczych obok nauk, które badają aktualne struktury świata i stąd mogą (jeśli inne względy nie stoją na przeszkodzie) z nimi eksperymentować, takich jak mechanika nieba, mineralogia, biologia funkcjonalna, istnieją także nauki badające pochodzenie tych struktur, takie jak kosmogonia, geologia historyczna, biologia ewolucyjna. Historia Ziemi, historia życia na Ziemi i historia ludzkości rozumiane jako *res gestae* są trzema fragmentami tego samego procesu. Rozumiane jako *historiae rerum gestarum* są trzema naukami znajdującymi się pod interesującym nas względem w tym samym położeniu, co i historia jednostek i ludzkości. Nędzna to jednak pociecha.

Ale historia hamuje zapędy eksperymentalne jeszcze w inny sposób. Rzeczywistość społeczną czy psychiczną współwyznaczają przeszłe zdarzenia. Aktualne w danej chwili stany systemów społecznych i osobowości są bardziej produktami historii i biografii niż swobodnego *fiat* współczesnych. Zatem, takie a nie inne podłoże historyczne i uformowane wcześniej struktury psychiczne wybitnie ograniczają ilość i wielkość zmian, jakie można w badanych systemach wprowadzić. Wachlarz tych zmian jest szeroki, ale nie nieograniczony, nie można od historii oderwać się zupełnie i nadawać zmiennym społecznym wartości dowolnych, nie można też tworzyć ich dowolnych kombinacji. Stan wyjściowy systemów społecznych odziedziczony po przeszłości zazwyczaj musi być traktowany jako dany. „Ludzie sami tworzą swoją historię — pisał Marks — ale nie tworzą jej dowolnie, nie w wybranych przez siebie okolicznościach, lecz w takich, jakie już zastali, w okolicznościach danych i przekazanych. Tradycja wszystkich zmarłych pokoleń ciąży jak zmora na umysłach żyjących”[106]. Tyle myślenie naukowe.

Niedocenianie ograniczeń zmienności zjawisk społecznych dyktowanych przez dziedzictwo historyczne właściwe jest za to pewnym odmianom myślenia utopijnego. Wyraziło się ono w micie „totalnej innowacji” — „zaczynania [życia społecznego] od nowa” oraz w „micie rewolucji” jako „całkowitego zerwania historycznej ciągłości, jako jednorazowego przeskoku , do doskonałości, jako błyskawicznej kreacji społeczeństwa idealnego”[107]. W mitologiach tych obecna jest wiara w możliwość zupełnego uwolnienia się od historii.

Dążenie do zerwania z obciążeniami historii miało charakter programowy w społecznościach utopijnych wieku XIX. Badacz zjawisk społecznych- może dostrzec w ich tworzeniu pewien rodzaj eksperymentowania społecznego[108]. Komuny były zakładane po to, by z dala od odrzuconego społeczeństwa osiągnąć życie zgodne z ideałem: zlikwidować wynaturzenia prywatnej własności i systemu fabrycznego a przez życie w harmonii i właściwą edukację wychować nowego człowieka. Jednak ideologowie i założyciele wspólnot utopijnych kierowali się również motywacjami poznawczymi: chcieli dowieść sceptycznej większości, że nowe formy organizacji społecznej są możliwe do realizacji i wykazać praktycznie ich przewagę nad formami panującymi.

„W swym najbardziej podstawowym sensie socjalizm utopijny pierwszej połowy XIX w. był rewolucją przeciw semifatalizmowi, który przedstawiał społeczeństwo jako produkt praw naturalnych, oskarżał proletariat o to, że jest jedynym sprawcą własnej nędzy, i ostro zaprzeczał możliwości «sztucznego» polepszenia warunków poprzez konstruktywne ustawodawstwo. Socjalizm utopijny zaprzeczał tym założeniom [i głosił idee], że ludzki umysł i pomysłowość mogą sprostać zadaniu zbudowania nowego porządku społecznego. Rzecznicy tego kierunku twierdzili, że natura ludzka jest przede wszystkim produktem środowiska społecznego,

[106] K. Marks, 18 brumaire’a Ludwika Bonaparte, w: K. Marks. F. Engels, Dzieła, t. 8, Warszawa 1962, s. 121.

[107] J. Szacki, Utopie, Warszawa 1968, s. 150 - 168 oraz Tradycja, Warszawa 197], s. 25-31.

[108] Zob. R. M. Kanter, Commitment and community. Communes and utopias in sociological perspective, Cambridge, Mass. 1972.

a lekarstwo na istniejące zło można znaleźć tworząc lepsze instytucje społeczne”[109].

Tworzone przez socjalistów utopijnych wspólnoty były pomyślane jako praktyczna demonstracja słuszności tej filozofii społecznej. I tak np. Robert Owen zakładając swą Nową Harmonię chciał dowieść, że możliwe jest zbudowanie społeczeństwa bez własności prywatnej, religii i współczesnej formy małżeństwa — trzech wielkich przeszkód zagradzających mu drogę do przebudowy społeczeństwa.

Wspólnoty te pomyślane były również jako początek działania w skali całego społeczeństwa. Ich założyciele widzieli w nich wzory przyszłego społeczeństwa doskonałego, oczekiwali, że społeczeństwo oświecone przykładem pójdzie śladem wspólnot. Tak np. Charles Fourier widział odrodzone społeczeństwo jako światową federację falansterów. Wzgląd ten upodabnia zakładanie wspólnot do eksperymentalnego pilotażu w małej skali zmian społecznych planowanych na szerszą skalę.

We wspólnotach utopijnych oderwanie się od historii było traktowane jako warunek wstępny zbudowania nowej społeczności lub wypracowania rozwiązań najskuteczniejszych. Wierność tradycji nie dałaby się pogodzić z tymi dążeniami. Stąd w jednej z takich wspólnot istniał nawet zakaz studiowania historii, by nie wprowadzać do miłującej pokój komuny tonów wojennych i konfliktowych. Systematyczne badania Kanter nad dziewiętnastowiecznymi wspólnotami amerykańskimi dowodzą jednak, że komuny bez żadnych tradycji okazywały się mniej trwałe niż komuny zakładane na bazie zrzeszeń już istniejących.

Ciekawe, że klęski przedsięwzięć utopijnych entuzjaści eksperymentów ze wspólnotami brali nie za dowód fałszywości ich koncepcji, lecz za skutek działania „złych” tradycji. Tak np. Robert Owen nie przyznał po upadku swej Nowej Harmonii, że społeczeństwo bez własności prywatnej, małżeństwa monogamicznego i religii jest niemożliwe. Utrzymywał natomiast, że nie można zbudować takiego społeczeństwa z ludźmi obarczonymi „przesądami” wyniesionymi z odrzucanego społeczeństwa indywidualistycznego[110] zareagował on tak samo, jak niektórzy uczeni reagują na wyniki doświadczeń zaprzeczające drogim ich sercu hipotezom. Uczeni ci nie mówią wówczas, że widocznie ich hipotezy były błędne, lecz że nie zostały spełnione warunki kontroli eksperymentalnej.

Wiara w możliwość całkowitego i natychmiastowego uwolnienia się od historii może zmiany pobudzać, o ich powodzeniu decyduje jednak znajomość rządzących nimi mechanizmów, za którą postępuje świadomość ich granic.

8.2. KONTEKST STRUKTURALNY

Historyczny charakter zjawisk społecznych nie jest jedyną techniczną barierą eksperymentu — nie wszystkie zmiany możliwe w danych warunkach udaje się przeprowadzić. Niektóre granice zmiany społecznej są tak oczywiste, że aż niezauważalne. Inne wyraźnie dostrzega rozum potoczny i nauka. Są wreszcie i takie, o obecności których przekonują dopiero próby ich przekroczenia: o działaniach takich mówi się potem jak o „woluntaryzmie” czy nawet „awanturnictwie”, które nie liczyło się z „obiektywnymi prawami”.

Ograniczenie wspomnianych zmian dyktuje kontekst strukturalny zjawisk, które miałoby się ochotę zmieniać. Wzajemne powiązania między zjawiskami społecznymi, uwikłanie ich w szersze wysoko zorganizowane struktury, zmniejsza możliwości zmian skutecznych. Co więcej, bez rozeznania się w tych związkach nawet nie sposób określić stopnia manipulacji zjawiskami interesującymi badacza jako

[109] H. Becker, H. E. Barnes, Rozwój myśli społecznej od wiedzy ludowej do socjologii, Warszawa 1965, t. 2, s. 253.
[110] Zob. F.S. Chapin, Experimental designs in sociological research, New York 1955, s. 9.

przedmioty ewentualnych oddziaływań. Systemy społeczne składają się bowiem z elementów powiązanych ze sobą w różnym stopniu — istnieje tu daleko posunięte zróżnicowanie zarówno między systemami, jak i między ich elementami. Kontekst strukturalny, w jakim występują wyróżnione zjawiska społeczne powoduje dwie ważne konsekwencje.

Manipulacja zmierzająca do wywołania zmian, których następstwa chce się lub chciałoby się określić, odbywać się musi za pośrednictwem wielu czynników. Jest ona szczególnie trudna wtedy, gdy czynniki te mają charakter nie addytywny, lecz interakcyjny. Manipulacja musi wówczas obejmować wiele czynników na raz. Jeśli, co gorsza, manipulowalność ich jest niejednakowa, to możliwości eksperymentatora określane są dostępnością elementu najmniej manipulowalnego. Jest to sytuacja próby zmiany podsystemu, którego działanie zależy od systemu nadrzędnego, jednakże bez zmiany tego drugiego.

Równie poważne problemy nastręcza kontrola eksperymentalna. Nie można dowolnie oddzielać jednych czynników od drugich i zmieniać jednych bez równoczesnej zmiany drugich. Stwarza to trudności dla kontroli przez tworzenie układów względnie izolowanych. Nie można również tworzyć dowolnych układów kontrolnych — oznacza to trudność kontroli za pomocą sytuacji porównawczych.

I znowu, podobnie jak w przypadku kontekstu historycznego, niedocenianie siły kontekstu strukturalnego jest właściwością myślenia utopijnego. Towarzyszy mu przekonanie o możliwości dokonywania , zmian cząstkowych bez równoczesnej przebudowy całego społeczeństwa, zmiany osobowości bez naruszania kształtujących ją warunków społecznych, przebudowy życia społecznego wedle woli i stosownie do idei społeczeństwa doskonałego. Podejmowane przez różne ruchy społeczne od socjalizmu utopijnego do hippisów próby organizowania komun wewnątrz społeczeństw opartych na zasadach komunom przeciwstawnych, to praktyczne konsekwencje takich przekonań. Testem empirycznym tych przekonań są losy komun. Wynik testu jest negatywny, przytłaczająca część owych społeczności miała żywot motyli.

Jednakże badania nad wspólnotami dziewiętnastowiecznymi i współczesnymi rzucają nader interesujące światło na trudności tworzenia pewnego typu społeczności eksperymentalnych[111]. Wiele z tych komun dążyło do izolowania się od szerszego środowiska społecznego, tak samo jak i od tradycji. Izolacja może mieć charakter przestrzenny lub tylko aksjologiczny. Z izolacją przestrzenną mamy do czynienia wtedy, gdy wspólnota organizuje się z dala od skupisk ludzkich, gdy kontroluje ona kontakty swych członków z otoczeniem, gdy jej członkowie zaspokajają wszystkie swe potrzeby w obrębie wspólnoty. Izolacja aksjologiczna polega na zdecydowanym wyborze przez członków pewnego stylu życia i równie kategorycznym odrzuceniu stylów konkurencyjnych. W tym drugim wypadku bariera geograficzna nie jest konieczna, aczkolwiek pomaga ona członkom wspólnoty wytrwać przy podjętym wyborze i przyczynia się do ukształtowania odrębnej kultury zbiorowości.

Izolacja przestrzenna nie stanowiła problemu w społeczeństwie dziewiętnastowiecznym, gdy niewielkie potrzeby członków można było zaspokoić w ramach jednej zbiorowości i autarkicznej gospodarki; społeczeństwo było stabilne, a przemieszczanie się ludzi, towarów' i idei nieznaczne. W ówczesnym społeczeństwie homogenicznym kulturowo przy niewielu alternatywnych stylach życia łatwiej było wybrać jeden z nich i łatwiej było znaleźć ludzi, którzy dokonaliby tego samego wyboru. W społeczeństwach współczesnych — zurbanizowanych, ruchliwych, rozwiniętych technologicznie, odznaczających się pluralistyczną kulturą i rozwiniętymi środkami jej przekazu, izolacja przestrzenna wspólnoty jest niemożliwa i niewystarczająca dla kontroli życia zbiorowości. Konieczna jest izolacja aksjologiczna — w kulturze heterogenicznej i ona jest trudniejsza niż kiedykolwiek.

W przeciwieństwie do myślenia utopijnego, myślenie naukowe skłania się raczej do przeceniania niż do niedoceniania siły powiązań strukturalnych wewnątrz

[111] Zob. Kanter, op. cit.

systemów społecznych. Prowadzić to może albo do negacji możliwości eksperymentu, albo do uznania za jedyny sensowny eksperyment eksperymentu w skali całego społeczeństwa; są to — rzecz jasna — stanowiska skrajne. Z właściwego historycznego przekonania o synchroniczności zjawisk społecznych — o ścisłej ich odpowiedniości w określonym miejscu i czasie[112] wynika teza, że „eksperyment społeczny (jeśli istnieje coś takiego) mógłby mieć wartość tylko wtedy, gdyby był przeprowadzony w skali holistycznej”[113]. Na jednym z tych stanowisk stoi organicyzm, na drugim marksizm, obydwa — mimo różnic — mieszczące się w nurcie historyzmu.

I tak Comte, który w społeczeństwie dostrzegał superorganizm, w którym wszystko ze wszystkim się wiąże, nie widział w instrumentarium socjologii miejsca na eksperyment. Zmieniając jeden czynnik, musielibyśmy zmienić zarazem wszystkie inne — argumentował — a przecież wszystkiego obserwować nie sposób, zatem eksperyment jest w socjologii bezużyteczny. Z kolei Spencer dostrzegający w społeczeństwie doskonalą harmonię twierdził, że każda zmiana eksperymentalna zakłócałaby równowagę społeczną, a że stan ten on cenił, więc konkludował, że eksperyment socjologiczny jest niebezpieczny[114].

Natomiast marksizm uszczegółowiający założenie o synchroniczności zjawisk społecznych w kierunku Uwydatniania „jedności walki i przeciwieństw” a nie organicznej harmonii doszedł w interesującej nas tu kwestii do odmiennych wniosków praktycznych. Rewolucja zastępująca jedno społeczeństwo drugim, a nie reformistyczne i utopijne zmiany cząstkowe to właśnie „eksperyment społeczny przeprowadzony w skali holistycznej”, tyle że dokonany nie z myślą sprawdzenia teorii rewolucji, ale z zamiarem bezpośredniej realizacji projektu ideologicznego. Świadome kierowanie historią w takiej skali ma za warunek powodzenia właśnie rozeznanie w czynnikach i mechanizmach zmiany, a nie sam tylko optymizm co do jej możliwości, co miało miejsce w myśli utopijnej.

8.3. OGRANICZONA WŁADZA BADACZY

Trzecią — po ograniczeniach historycznych i strukturalnych — z technicznych barier eksperymentowania na systemach społecznych i ludziach jest mały zakres władzy badacza nad interesującymi go zjawiskami. Nawet jeśli już w granicach wyznaczonych historycznym charakterem zjawisk i kontekstem strukturalnym możliwa jest manipulacja pewnymi czynnikami, to wcale stąd nie wynika, że możliwość ta otwarta jest dla badacza.

Manipulowalność zjawisk jest własnością relatywną. Najpierw w stosunku do zmiennych: atrybuty wielkich struktur społecznych są mniej podatne na zmiany niż własności małych grup. Następnie, co nas tu szczególnie interesuje, w stosunku do podmiotów działających: te same zjawiska społeczne są w różnym stopniu podatne na odkształcające oddziaływanie różnych ośrodków. Centra władzy politycznej i ekonomicznej dysponują znacznymi możliwościami manipulacji, nawet zmiennymi makrosocjologicznymi. Nawet dla nich struktura społeczna nie jest jednak doskonale plastyczna, a posiada pewną sztywność niepokonalną; możliwości zorganizowanych ruchów masowych, dyktatorów i innych decydentów najwyższego szczebla są również ograniczone. Tym bardziej więc świat społeczny nie poddaje się zabiegom badaczy — ludzi nie władzy, a wiedzy. Władza przypisana roli społecznej uczonego wystarcza tylko do manipulowania zmiennymi charakteryzującymi małe struktury społeczne; co więcej, nie grupy rzeczywiste, naturalne, lecz sztuczne, stworzone przez niego samego na czas i użytek eksperymentu.

[112] Zob. J. Szacki, O tzw. historyzm w naukach społecznych, w: S. Nowak (red.), Metodologiczne problemy teorii socjologicznych, Warszawa 1971.

[113] Jest to fragment rekonstrukcji historyzmu, dokonanej przez Poppera w The poverty of historicism, London 1957, cyt. ze s. 84.

[114] Zob. D. Martindale, Community, character, and civilisation. Glencoe 1963, s. 36.

Rozważmy dla przykładu twierdzenie, że „ruchliwość społeczna w górę lub w dół powoduje wzrost uprzedzeń etnicznych”[115], Gdyby odnieść je do zmian pozycji całych klas i warstw w strukturze społecznej, to badacze wyposażeni we władzę polityków mogliby je testować eksperymentalnie, lecz tylko *w* granicach autonomii polityki względem struktury społecznej. Gdyby natomiast twierdzenie to odnieść do mniejszych struktur, to tylko badacze nie mogliby go eksperymentalnie kontrolować: zmiana kierunku czy tempa ruchliwości mniejszych fragmentów struktury społecznej jest technicznie wykonalna, tyle że nie dla nich. Nie mają oni władzy wystarczającej nawet do dokonania zmian pozycji społecznej jednostek w małej, ale naturalnie istniejącej grupie, takiej jak rodzina czy grupa zadaniowa w przemyśle lub wojsku. To, co im pozostaje, to albo wykrywanie psychologicznych następstw procesów zachodzących w społeczeństwie bez ich wpływu, albo manipulacja pozycjami społecznymi członków sztucznie stworzonych grup eksperymentalnych. Rozważane twierdzenie jest wynikiem wykorzystania pierwszej’ z tych możliwości, ale potwierdzają je również wyniki drugiego rodzaju.

Naturalną konsekwencją tego stanu rzeczy jest fakt, że sami badacze nie dokonują eksperymentów na całościach większych niż mała grupa. Czynią to natomiast wespół z administratorami życia społecznego i instytucji publicznych — ci delegują im część swoich uprawnień, a wynikiem tego są opisywane już eksperymenty społeczne.

Swego rodzaju eksperymentami społecznymi są także, jak już pisaliśmy, wspólnoty utopijne. Ale znowu eksperymentatorami nie są tu badacze, lecz grupy lub ich charyzmatyczni przywódcy. Jeśli przedsięwzięcia takie możliwe są mimo braku władzy administracyjnej nad grupą, to dzieje się to dzięki silnemu zaangażowaniu członków grupy w jej wartości i ideały. I Znowu, badacz jako badacz, nosiciel pewnej roli społecznej, nie ma możliwości kontroli większych zbiorowości przy pomocy symboli i wartości.

8.4. ROZCIĄGŁOŚĆ ZJAWISK W CZASIE

Dodatkowym źródłem ograniczeń eksperymentalnego badania zjawisk społecznych jest ich rozciągłość w czasie, niewspółmierna z możliwym lub tylko typowym czasem trwania eksperymentu. Ten zamknięty jest dopuszczaną przez badanych długością i częstotliwością ich przebywania w laboratorium, dopuszczaną długością zmienionego funkcjonowania instytucji społecznych, trwałością grup eksperymentalnych, cierpliwością i stałością zainteresowań samych badaczy, pilnością zagadnień praktycznych, których rozstrzygnięciu eksperymenty miałyby służyć itp. Okres wyznaczony tymi czynnikami — jakkolwiek elastyczny, nie jest jednak dowolnie długi i nawet dla eksperymentów społecznych nie przekracza on kilku lat. Tymczasem do tego, aby wywołać niektóre zjawiska wyspecyfikowane w poprzednikach hipotez socjologicznych, potrzebny byłby — poza spełnieniem wszystkich warunków, o których była dotąd mowa — dosyć długi okres. Z tego punktu widzenia zjawiska społeczne można byłoby podzielić na trzy grupy. W pierwszej znalazłyby się te, które można byłoby Wywołać natychmiast, w trakcie jednego lub paru kolejnych seansów eksperymentalnych — są one dziedziną eksperymentowania socjopsychologicznego tak w laboratoriach jak i w środowiskach naturalnych. Do drugiej grupy weszłyby zjawiska, które trzeba byłoby wywoływać przez dłuższy, nawet kilkuletni okres — jest to teren eksperymentów społecznych, badających skuteczność pewnych reform społecznych; samo wprowadzenie takich reform wymaga znacznej ilości czasu. Wreszcie, grupa trzecia składałaby się ze zjawisk, których wywołanie, gdyby nawet było możliwe technicznie, wymagałoby wielu lat, jeśli już nie pokoleń. Przykładem może być*tu choćby ruchliwość spo-

[115] J. Greenblum, L. J. Pearlin, Ruchliwość pionowa a uprzedzenia etniczne, w: A. Malewski (red.), Zagadnienia psychologii społecznej, Warszawa 1962, s. 243.

łeczna międzypokoleniowa, mierzona stosunkiem zawodów synów do zajęć ojców. Toteż zjawiska tego rodzaju również i z dyskutowanego powodu nie nadają się na bodźce eksperymentalne operowanie nimi, nawet gdyby było możliwe, kłóciłoby się z naturalnym rytmem badań naukowych.

Na tym jednak rzecz się nie kończy. Rozciągłość czasowa procesów i zjawisk społecznych nie tylko utrudnia wprowadzenie i kontrolę licznych zmiennych określonych w poprzednikach hipotez socjologicznych, lecz także utrudnia obserwację wielu innych zmiennych figurujących w następnikach hipotez. Wynika to stąd, że odstęp czasowy między niektórymi zjawiskami pozostającymi ze sobą w związkach przyczynowych jest bardzo długi.

Tak dzieje się choćby w eksperymentalnych próbach innowacji społecznych. Eksperyment społeczny wymaga starannego zbadania dalekosiężnych skutków wprowadzonych zmian społecznych. Niestety, wymóg ten odbiera często sens robocie eksperymentalnej. Wartość pewnych propozycji praktycznych można ocenić dopiero po upływie wielu lat a wówczas może się okazać, iż warunki wyjściowe zmieniły się tak znacznie, że sprawdzone rozwiązania są już przestarzałe. Kryterium wartości podjętych w celach eksperymentalnych reform szkół wyższych nie są w gruncie rzeczy wyniki egzaminacyjne studentów, lecz poziom pracy absolwentów w ich nowych środowiskach społecznych i zawodowych — ocena wyników takich eksperymentów wymagałaby więc utrzymania kontaktu z absolwentami przez kilka przynajmniej lat po opuszczeniu przez nich murów uczelni. Zagadnienie to występuje jeszcze dobitniej w eksperymentach z zakresu polityki resocjalizacyjnej i penitencjarnej. Skuteczność praktyk resocjalizacyjnych mierzy się odsetkiem recydywistów wśród osób, które w swoim czasie poddane były tym praktykom — jest to więc miara wymagająca szczególnie długiego czasu. W tej sytuacji wnioski eksperymentalne tracą na pewności — dobrze sobie zdają z tego sprawę inicjatorzy niektórych eksperymentów.

Istnieją też takie procesy społeczne, których charakter długofalowy, jeśli już nie wręcz historyczny, zdecydowanie wyklucza możliwości ich eksperymentalnego badania. Nawet gdyby można było wywoływać je, to nie można byłoby ich obserwować, bo cierpliwości albo i życia by na to nie stało. Taki charakter mają procesy makrosocjologiczne, to do nich właśnie odnosi się powiedzenie *sero molunt deorum molae* — wolno mielą młyny bogów.

Zaznaczyć trzeba na koniec, że barierą eksperymentowania jest nie tylko zbytnia powolność, ale i zbytnia szybkość badanych procesów utrudniająca ich szczegółową obserwację. Taki charakter mają niektóre procesy uczenia się u ludzi, toteż woli się badać je u zwierząt, gdzie występują w tempie zwolnionym.

*

* *

Gdyby chcieć określić w sposób ogólny i zwięzły zasięg zjawisk objętych wskazanymi w tej części barierami eksperymentowania, to można by rzec, że siła tych barier rośnie w miarę zwiększania się rozmiarów i komplikacji zjawisk. Inaczej — w miarę przesuwania się z poziomu małych grup na poziom wielkich struktur społecznych. Procesy i zjawiska występujące na poziomie pierwszym są zawsze łatwiej manipulowalne, z zasady są też prostsze i szybsze (co więcej, można je tworzyć i reprodukować do woli) niż procesy występujące na drugim poziomie.

Różnice w manipulowalności zjawisk makro- i mikrospołecznych pociągają też za sobą — co zrozumiałe — istotne zróżnicowania między makro- a mikro- socjologią. Pierwsza jest nauką nieeksperymentalną w stopniu nie mniejszym niż astronomia, uważana za wzorzec dyscypliny nieeksperymentalnej. Nie tworzy ona przez stymulację i kontrolę sztucznych sytuacji spełniających wymogi indukcji, ale usiłuje je znaleźć w minionej i współczesnej strukturze i dynamice społecznej. Za to mikrosocjologia oparta jest na eksperymencie w stopniu poważnym. W konsekwencji, dzięki zaletom eksperymentu mikrosocjologia jest dyscypliną bardziej rozwiniętą od socjologii wielkich struktur. Podział na makro- i mikrosocjologię ma tedy za podstawę taki podział przedmiotu badania, z którym skorelowane są moż-

liwości manipulacji zjawiskami badanymi a w konsekwencji — również obfitość i pewność zgromadzonej wiedzy.

Techniczne ograniczenia możliwości eksperymentowania nie są jednymi — nie mniej istotne są bariery aksjologiczne.

9. AKSJOLOGICZNE GRANICE MANIPULACJI EKSPERYMENTALNYCH

Jedną z osobliwości nauk o człowieku tak społecznych jak i biomedycznych jest fakt, że wartość zabiegów badawczych w nich stosowanych jest zagadnieniem nie tylko metodologicznym, lecz i aksjologicznym: można je oceniać nie tylko w świetle standardów poznawczych, ale i uwikłanych w badanie wartości pozapoznawczych, takich jak życie i zdrowie ludzkie, trwałość ładu społecznego, status społeczny, integralność osobowości, autonomia woli, godność, prywatność, pozytywna samoocena itp. Istnieje więc możliwość, że nie wszystko, co służyłoby gromadzeniu wiedzy o ludziach jest dopuszczalne ze stanowiska panujących w społeczeństwie systemów wartości, wyrażonych w ideologiach, prawie i etyce. I tak jest faktycznie.

9.1. ZDERZENIE SIĘ WARTOŚCI

Nie wszystkie manipulacje eksperymentalne naukowo interesujące a technicznie wykonalne są dozwolone przez panujące systemy wartości społecznych — zakres eksperymentów możliwych ulega przez to dalszemu zawężeniu. To tylko fikcyjny Idealny Eksperymentator chciałby i zarazem byłoby mu wolno poddawać rozmaite fragmenty życia społecznego i osobniczego wymaganiom schematów eksperymentalnych, izolować jedne od drugich, wyrywać zjawiska z jednych warunków a umieszczać w drugich, poddawać społeczeństwo interesującym z naukowego punktu widzenia manipulacjom itp., a wszystko to po to tylko, by sprawdzić pewne hipotezy naukowe. Jego — również fikcyjną — ideologią byłby rozumiany dosłownie a więc sprowadzony do absurdu scjentyzm, albo- ’ wiem „...ideologia scjentystyczna jest taką ideologią, która postuluje, by społeczeństwo w całości nastawione było na realizację tych samych zasad, do których dąży nauka”[116].

Program taki popadałby jednak w jawny konflikt ze wskazanymi wyżej wartościami ludzkimi i społecznymi. Ich dominacja stanowi właśnie granice eksperymentowania na ludziach i systemach społecznych. Obfitość dobrze uzasadnionej wiedzy, choć niezwykle cenna, nie jest przecież ani jedyną ani najważniejszą wartością w naszej kulturze i nie ona jest preferowana w każdym przypadku konfliktu dążeń poznawczych z pozapoznawczymi wartościami ludzkimi i społecznymi.

Zakazy strzegące wartości ludzkich są na ogół silne i dobrze zinternalizowane. Niektóre z nich są wręcz tak oczywiste, że mało komu przychodzi na myśl ich przekraczanie, podobnie jak i konieczności przyrodniczych, choćby zasady ciążenia. Ale owa „naturalność” utrudnia również postronnemu obserwatorowi ich dostrzeżenie — łatwo natomiast zauważyć je na porównawczym materiale historycznym, w procesach ich powstawania, obumierania i łamania. I tak:

Eksperymenty starożytnych władców Egiptu, na które się wyżej powołaliśmy, mniejsza o to, czy faktycznie przeprowadzane, możliwe były lub byłyby tylko w społeczeństwach, w których jeszcze nie wykształciły się idee świętości życia

[116] S. Amsterdamski, Nauka współczesna a wartości, „Zagadnienia Naukoznawstwa” 1971, nr 1.

i autonomii jednostki wobec władcy i państwa. Z kolei upowszechnienie się religijnej koncepcji człowieka jako nienaruszalnej całości danej od Boga uniemożliwiało jakiekolwiek doświadczenia z ludźmi. Eksperymentalna nauka o człowieku stała się możliwa dopiero po odejściu od tej koncepcji. Siedemnastowieczny mechanizm, dostrzegający w ludziach jedynie układy rurek, sznurków i przekładni, nie wykluczał już eksperymentu z instrumentarium tej nauki[117].

Bariery te widoczne są jednak szczególnie wyraźnie w sytuacjach ich łamania. Myślę tu o doświadczeniach dokonywanych w hitlerowskich obozach koncentracyjnych. (Były to wprawdzie eksperymenty na organizmach raczej niż na psychice, z punktu widzenia celów, dla których nad nimi się zatrzymujemy, okoliczność ta jest wszakże bez znaczenia). We wspomnianych doświadczeniach usiłowano określać minimalny poziom pożywienia, opieki lekarskiej czy temperatury, przy którym człowiek może jeszcze utrzymać się przy życiu, a nawet ciężko pracować. Dokonywano na ludziach prób środków farmakologicznych i operacji medycznych, które przedtem i potem przeprowadzano na zwierzętach, czasami nawet nie bez etycznych niepokojów, jak wskazuje *casus* dyskusji nad wiwisekcją. Drastyczność sytuacji i wyraźny charakter jej społecznych uwarunkowań znakomicie ułatwiają jej obserwację.

Interesującą jest radykalna opinia, że „Wśród wielu «eksperymentów» prowadzonych pod płaszczykiem nauki w czasach Hitlera nie ma ani jednego, który mógłby być poważnie rozważany na łamach czasopisma naukowego. Sadyzmu, z którym «eksperymenty» były prowadzone, nie okupiły one żadnym wartościowym wynikiem naukowym”[118]. Stało się tak zapewne z powodu małego udziału myślenia teoretycznego w tych doświadczeniach a podporządkowaniu ich pytaniom doraźnym, przy czym potrzeby praktyczne, które im dały początek, są zbyt zbrodnicze i nietypowe, aby wiedza im służąca znajdowała szersze zastosowanie. Niezależnie od tego, nie ma wyników naukowych, które mogłyby wynagrodzić poniesione straty moralne: nawet gdyby eksperymenty te dały jakieś niezwykle cenne i wartościowe w skutkach wyniki, to i tak budziłyby kategoryczny sprzeciw wypływający z trwałych i uniwersalnych wartości naszej kultury.

Trzy z nich są najważniejsze. Pierwszą jest uznanie życia ludzkiego za dobro najwyższe, którego nie wolno traktować instrumentalnie, lecz tylko autoteliczne. Drugą — autonomia woli, prawo do decydowania o własnym losie. Trzecią jest unikalność osoby ludzkiej, uznanie jej za byt jedyny w swoim rodzaju i niepowtarzalny[119].

Paradoksalne, ale sytuacja podeptania podstawowych wartości ludzkich ukazuje szczególnie wyraźnie podporządkowanie motywów poznawczych względom praktycznym. Sami eksperymentatorzy mogli wprawdzie subiektywnie kierować się motywacjami poznawczymi, jednak dysponenci władzy dawali im prawo czynienia zbrodniczych doświadczeń bynajmniej nie po to, by zgodnie z ideałami niemieckich uniwersytetów *treiben Wissenschaft,* lecz po to by ułatwić ekspansję jednych a eksterminację drugich.

Współczesna postać i zakres eksperymentalistyki w naukach społecznych a i ogólniej — w naukach o człowieku, jest w zasadzie wynikiem podporządkowania wartości poznawczych wartościom ludzkim. Tak więc zakres eksperymentowania jest niewielki w porównaniu z możliwościami technicznymi. Istnieją bowiem trzy[1] bariery, których przekroczenie jest co prawda wykonalne technicznie, ale moral-

[117] Zob. H. Hörz, The relation of modern scientific conceptions to the human image, w: P. Weingartner, G. Zecha (eds.), Induction, physics, and ethics, Dordrecht 1970.

[118] E. L. Kelly, J. C. Lingoes, Przetwarzanie danych w badaniach psychologicznych, w:_H. Borko (red), Maszyny cyfrowe w badaniach naukowych, Warszawa 1969, s. 69.

[119] Materiał badawczy „[…] był zawsze towarem, produktem masowym, którego każda jednostka była dowolnie wymienialna na jakąś inną” — pisze Bettelheim. Autor ten cytuje ustęp z korespondencji między laboratorium farmakologicznym zakładów IG Farben a komendanturą obozu w Oświęcimiu dotyczącej warunków kupna-sprzedaży ludzi potrzebnych do doświadczeń. Istnieją tam tylko kwestie ceny, ilości i stanu, w jakim mają oni być dostarczeni. Gdyby z korespondencji tej usunąć słowo „kobieta”, nikt nie domyśliłby się, że idzie w niej nie o rzeczy, lecz o osoby. Zob. B. Bettelheim, The informed heart, London 1970, 6. 221 - 5.

nie czy prawnie zakazane lub obciążone dwuznacznością etyczną. Owe bariery to: naganność samych manipulacji eksperymentalnych, szkodliwość ich zamierzonych i ubocznych konsekwencji oraz etycznie wątpliwy charakter samego manipulowania ludźmi bez względu na następstwa. Rozpatrzymy je rozpoczynając od ideologiczno-politycznych, a następnie przejdziemy do etycznych ograniczeń eksperymentowania na zbiorowościach i osobnikach ludzkich.

9.2. IDEOLOGICZNE GRANICE EKSPERYMENTU

Eksperyment jest pewnym rodzajem zmiany społecznej, a zatem ideologiczne oraz polityczne bodźce i hamulce zmiany działać będą również w przypadku eksperymentu. Ideologia i polityka dopuszczają tylko takie zmiany społeczne, które nie są sprzeczne z ich celami, tj. utrzymaniem bądź kierunkową zmianą porządku społecznego. Będą to więc zmiany dotyczące czynników mniej istotnych niż warunki utrzymania władzy oraz równowagi czy rozwoju systemów społecznych. Co więcej, ich aprobatę uzyskują tylko zmiany o zasięgu ograniczonym, takie, których skutki są kontrolowane i odwracalne.

Doskonałą ilustracją tych tez jest referowany już eksperyment nad skutecznością emocjonalnej i racjonalnej propagandy wyborczej wykonany w 1935 r. przez Hartmanna w jednym z miast Stanów Zjednoczonych[120]. Siły głównych partii — demokratycznej i republikańskiej były tak dokładnie wyrównane w tym mieście, że żadna kampania eksperymentalna na rzecz którejkolwiek z nich nie była możliwa ze względu na zagrożenie pozycji drugiej partii. Natomiast popularność partii socjalistycznej była tak niewielka, że jej zwycięstwo nie wchodziło w rachubę; Hartmann mógł więc przeprowadzić eksperymentalną kampanię propagandową na rzecz partii socjalistycznej — nie miała ona wpływu na zmianę układu sił politycznych w mieście.

Eksperymenty dotyczące spraw, o których mowa, mogłyby przybrać postać eksperymentów społecznych w sensie ograniczonych prób zmiany społecznej. Nie we wszystkich porządkach społecznych istnieją warunki do tego rodzaju przedsięwzięć i nie wszystkie ideologie w jednakowym stopniu dopuszczałyby innowacje eksperymentalne. Podstawowa różnica istnieje tu między porządkami i ideologiami konserwatywnymi a porządkami i ideologiami radykalnymi.

„Sednem radykalizmu jest przekonanie, że stan rzeczy może i powinien być zmieniony. Ponadto zmiany te mają iść w kierunku nowym i niewypróbowanym. Jest to więc śmiała wyprawa na niezbadany ocean życia”[121]. Sam więc charakter myślenia radykalnego może zachęcać pewne porządki społeczne do posługiwania się eksperymentem jako narzędziem czyniącym politykę społeczną polityką racjonalną, pozwalającym oceniać szanse przedsięwzięć jeszcze przed wprowadzeniem ich w życie.

Z narzędzia tego nie mogą natomiast czynić użytku porządki konserwatywne. Koncepcja społeczeństwa jako harmonijnego organizmu i apologia równowagi, właściwe myśleniu konserwatywnemu[122], prowadzą do uznania eksperymentu albo za niemożliwy albo za szkodliwy. Myśl konserwatywna oscyluje, jak zauważa filozof „między przeświadczeniem, że świat jest tak doskonały, że nie wymaga naprawy a przeświadczeniem wedle którego świat jest tak marny, że naprawić go niepodobna”. Prostą konsekwencją tego wyznania wiary, w obu jego odmianach, jest hasło „żadnych eksperymentów!” — z dawien dawna ulubione zawołanie konserwy społecznej. Mowa w nim o eksperymentach w sensie śmiałych innowacji

[120] G. W. Hartmann, A field experiment on the comparative effectiveness of „emotional” and „rational” political leaflets in determining election results, w: R. L. Rosnow, E. J. Robinson (eds.), Experiments in persuasion, New York 1967.
[121] P, M. George, Modele a konserwatyzm społeczny w socjologii, „Studia Socjologiczne” 1971, nr 2.
[122] Zob. K. Mannheim, Essays on sociology and social psychology, New York 1953, zwł. s. 103 - 111.

społecznych, ale gdy się jest przeciw zmianom, musi się być i przeciw ich eksperymentalnym próbom.

Na sprawę ideologicznych i politycznych uwarunkowań eksperymentowania w społeczeństwie można spojrzeć jeszcze z innej perspektywy.

Otóż, konserwatyzm posługuje się obok argumentów organicystycznych uzasadnieniami tradycjonalistycznymi. Wartość instytucji istniejących uzasadnia on przez odwołanie się do pozytywnie wartościowanej dawności, do tego, że wytrzymały one próbę doświadczenia historycznego, że wyrażają mądrość poprzedników, a że wszystko, co istnieje jest produktem historii, to i „wszystko, co istnieje, jest rozumne". Konserwatyzm wyklucza natomiast odwoływanie się do faktów i rozumu, odrzuca poszukiwanie uzasadnień merytorycznych[123].

Tak jak tradycjonalizm jest cechą konserwatyzmu, tak antytradycjonalizm jest właściwością ideologii i ruchów radykalnie postępowych. Te odrzucają wszelkie autorytety zewnętrzne, *a więc* i argumenty dawności czy samego istnienia, żądają uzasadnień merytorycznych — empirycznych czy rozumowych. Wierność tradycji jakiejkolwiek zasadniczo nie da się pogodzić z dążeniem do wypracowania rozwiązań maksymalnie zgodnych z przyjętymi wartościami, np. efektywnością. Tradycja co najwyżej wykluczałaby spod jurysdykcji empirycznej pewne rozwiązania — może ona więc tylko dostarczać propozycje, których dawność sama przez się w żadnym wypadku nie czyni ich wartościowymi. Instancją właściwą dla rozstrzygania sporów o wartość instytucji społecznych jest praktyka a narzędziem wypracowywania rozwiązań najlepszych — eksperyment.

O ile tradycjonalizm jest ideologią pewnych systemów społecznych, o tyle konsekwentny eksperymentalizm podniesiony do rangi ideologii społeczeństwa czyniącego efektywność swą wartością najwyższą, daje się tylko pomyśleć. Jego realizacją jest tylko Walden Two, fikcyjna społeczność wymyślona przez Skin- nera z głośnej powieści pod tym tytułem[124] W Walden Two wszystko, co istniało miało za sobą sankcję kontrolowanego eksperymentu, a argumentem w sporach o wartość konkurencyjnych urządzeń społecznych były zawsze protokoły eksperymentów przechowywane w archiwum komuny. Sposób parzenia herbaty, forma rodziny, system wychowania, kształt władzy politycznej — wszystko to było eksperymentalnie sprawdzone i uzasadnione. Efekty takiego postępowania dla rozwoju komuny i jej członków były nadzwyczajne. Okazało się przy tym, jak wielkie rezerwy tkwią w ludziach i w organizacji społeczności — eksperyment ujawniał je i przekonywał o ich wartości.

Konsekwentny antytradycjonalizm — odwrotna strona światopoglądu eksperymentalnego — był w Walden Two uderzający. Wyraża go dobrze *alter ego* Skinnera — ideolog i twórca komuny Frazier. Mówi on: „Wychowanie moralne należy do społeczności. Co do technik, to korzystaliśmy z każdej sugestii, jaką mogliśmy znaleźć, bez uprzedzeń wobec źródła. Ale nie przyjmowaliśmy jej na wiarę. Odrzucamy wszelkie prawdy objawione i poddajemy każdą zasadę testowi eksperymentalnemu. Aha, całkiem opacznie objaśniłem ci cały system, jeśli przypuszczasz, że którekolwiek z opisanych technik są niewzruszone. Poddajemy próbie wiele różnych technik i stopniowo dochodzimy do osiągnięcia zestawu najlepszego z możliwych. Nie przywiązujemy większej wagi do pozornego sukcesu zasady w historii. W Walden Two historię uznaje się tylko jako rozrywkę. Nie bierze się jej poważnie jako pokarmu dla myśli"[125].

Owa tak pięknie opisana postawa eksperymentalna występuje także w dawniejszych i współczesnych komunach, o których już wspominaliśmy.

Postawa ta wystąpiła szczególnie wyraźnie w Twin Oaks, społeczności skonstruowanej wg zasad technologii behawioralnej na wzór Skinnerowskiej Walden Two. Tak np. forma władzy została w Twin Oaks wypracowana eksperymentalnie. Zaczęto od anarchii, ta jednak nie dała oczekiwanych wyników. Skończono (1970)

[123] Zob. J. Szacki, Tradycja, Warszawa 1971, s. 31 -37.
[124] B. F. Skinner, Walden Two (1948), New York 1976.
[125] Ibidem, s. 115.

na utworzeniu zespołu planistów, którego skład odnawia się w ten sposób, że planiści ustępujący wyznaczają nowych a społeczność akceptuje ich z prawem weta. W zakresie stosunków interpersonalnych stosowano zrazu wzajemną krytykę, lecz potem okazało się, że brak całkowitej szczerości czyni ją trudną, poprzestano więc na krytyce tylko w parach[126].

Eksperymentalizm jako ideologia nie wyszedł jednak poza małe społeczności, z których ekstrapolacje mają wartość ograniczoną do pewnych tylko dziedzin, częściej np. stosunków międzyludzkich niż władzy politycznej. Permanentne eksperymentowanie nie jest, poza wszystkim innym, o czym już mówiliśmy i o czym mówić będziemy, możliwe z innego jeszcze powodu. Po pierwsze, w systemach społecznych obok dynamiki potrzebna jest pewna stabilność wielu urządzeń, choćby ze względu na ograniczone zdolności adaptacji jego członków do nowych praktyk. W działaniu społecznym potrzeba jest również trochę ślepej wiary w wartość obieranych sposobów postępowania. Nie sposób bowiem wątpić w dosłownie wszystko, coś trzeba przyjmować jako sprawdzone dostatecznie, radykalny sceptycyzm paraliżuje działanie.

Bariery polityczne i ideologiczne, wzbraniające eksperymentów, które mogłyby naruszyć założony stan równowagi lub kierunek rozwoju systemu społecznego oraz interesy zaangażowanych w ich wykorzystanie grup ludzkich, są podobnie jak i bariery techniczne tym wyraźniejsze, im większych czy istotniejszych grup dotyczyłyby manipulacje. Zmienne — atrybuty wielkich struktur i władzy, nie są w ogóle z tego punktu widzenia manipulowalne. Nie ma to jednak większego znaczenia, bo o ich niemanipulowalności przesądza już bariera techniczna. Zmienne — cechy instytucji mogłyby być manipulowane przez współpracujących ze sobą badaczy i administratorów; tu barierę techniczną można więc pokonać. To jednak, czy możliwość ta zostanie wykorzystana, zależy właśnie od wartości tych administratorów. Pozostaje jeszcze trzecia grupa zmiennych — własności osobników i małych grup. Ich eksperymentalne badania są technicznie możliwe. Nie wynika stąd jednak wcale, że są one niczym nie skrępowane, wyrasta bowiem przed nimi druga z barier aksjologicznych — bariera etyczna.

9.3. ETYCZNE GRANICE EKSPERYMENTU

Etyka zabrania przede wszystkim eksperymentalnego sprawdzania hipotez, które orzekają o zjawiskach etycznie negatywnych w swych poprzednikach: zakazuje stosowania manipulacji doraźnie dolegliwych dla badanych oraz sprzecznych z pewnymi wzorami funkcjonowania osobowości i grup ludzkich. Klasa ta obejmuje liczne hipotezy, które zaliczyłoby się do teorii nagród i (szczególnie!) kar, frustracji i agresji, konfliktów motywacyjnych, patologii społecznej. Chcąc testować je eksperymentalnie trzeba byłoby jednak czynić z ludzi męczenników nauki: karać ich, frustrować, doprowadzać do nędzy, obniżać ludziom pozycję społeczną i samoocenę, dezorganizować osobowości i systemy społeczne, wszystko to „dla dobra nauki” — po to tylko, aby sprawdzić, jakie będą reakcje na aplikowane bodźce. Choć wiedza w taki sposób uzyskana mogłaby być źródłem poważnych i trwałych korzyści dla społeczeństwa, to naganność takich poczynań jest zbyt oczywista, aby łatwo było o ich przykłady. Co najwyżej, spotyka się badania, w których manipulacja eksperymentalna jest tak blisko tego, co zakazane, że rodzą się wątpliwości, czy aby wszystko w tych eksperymentach jest w etycznym porządku.

Oto Bergin testował hipotezę o zależności rozmiaru zmiany w samoocenie pod wpływem nowych informacji od stopnia zaufania do informatora[127]. W tym celu podawał młodym ludziom m.in. fałszywe dane zaniżające ich pozycje na skali

[126] Zob. Twin Oaks — on to Walden Two, „Time” z 20.08. 1971; R. M. Kanter, Commitment and community, Cambridge, Mass. 1972, zwł. s. 18-31.

[127] A. E. Bergin, The effect of dissonant persuasive communications upon changes in self-referring attitude, „Journal of Personality” 1962, vol. 30, nr 2.

męskości i kobiecości w ten sposób, że brali je za faktyczne wyniki testów. Mimo ujawnienia fałszywości informacji po zakończeniu badania, manipulacja taka, pomijając już jej ewentualne trwalsze następstwa, prowokuje do postawienia pytania, czy nie jest ona dotkliwa dla badanych a naganna ze stanowiska etyki, która wartością czyni także pozytywną samoocenę, zwłaszcza jeśli ta jest zarazem adekwatna.

Dennis powtórzył z całą świadomością historycznych paranteli opisane wyżej doświadczenie faraona Psammetycha[128], Po to, aby określić, jakie reakcje dzieci są auto- a jakie socjogenne, zredukował opiekę nad parą niemowląt do minimum tj. do zapewnienia im wygody i dobrego samopoczucia. Tak więc izolował je na okres roku od wszelkich zbędnych kontaktów, nie warunkował w żaden sposób ani nie dostarczał okazji do aktywności. Choć eksperyment nie zdeformował rozwoju tych dzieci, to jednak może zrodzić się etyczny niepokój, czy aby owa „nienormalna" sytuacja badawcza nie naruszała zakazów zakłócania „normalnego" toku rozwoju w celach innych niż dobro ludzkie.

Etyka zabrania również ekspermentalnych testów hipotez orzekających o zjawiskach negatywnych w następnikach, inaczej mówiąc, każe liczyć się z możliwością szkodliwych następstw manipulacji eksperymentalnych. Oznacza to oczywiście zakaz takich manipulacji, z którymi sama hipoteza wiąże skutki niepożądane. Toteż takich eksperymentów się nie prowadzi — byłyby one wszak sadyzmem. Ale rozważany zakaz obejmuje też manipulacje, co do których brak pewności, że nie spowodują negatywnych zmian nie przewidzianych ani przez hipotezę, ani przez eksperymentatora. Problemem faktycznym w praktyce eksperymentalnej są dopiero manipulacje drugiego rodzaju. Trudno bowiem sądzić, że manipulacja zawsze wywołuje tylko zmiany, z którymi badacz liczył się i że jej następstwa znikają wraz z zakończeniem eksperymentu. Ich pojawienie się nieraz jest zaskoczeniem dla badacza i decyduje o odkrywczości badania, a wśród nich mogą znaleźć się fakty nie tylko obojętne, lecz i szkodliwe dla badanych, a z tymi eksperymentator liczyć się musi. O ten typ niebezpieczeństw ociera się wiele eksperymentów z zakresu nauk o zachowaniu.

Eksperyment Dennisa dowiódł m.in., że rozwój behawioralny izolowanych dzieci przebiegał identycznie jak w sytuacjach normalnych, zatem badanie go nie odkształciło. Ale był to dopiero wynik doświadczenia a nie jego przesłanka — badacz podejmował więc ryzyko nie jego osoby dotyczące.

Niezamierzonym następstwem niektórych eksperymentów psychospołecznych może być obniżenie się poczucia własnej wartości badanych. Niekiedy, jak w eksperymencie Bergina, dziać się może to za sprawą dostarczenia im fałszywych, zaniżających samoocenę informacji, których siły perswazyjnej być może nie jest w stanie pokonać poeksperymentalne uświadomienie badanym, że informacje te były fałszywe właśnie. W innych przypadkach sam schemat eksperymentu- zmierza do obnażenia badanym ich własnych słabości. Asch stawiał badanych wobec konieczności oceny pewnej bardzo prostej sytuacji w warunkach ewidentnie fałszywej, ale jednomyślnej opinii grupy[129] — znaczna część badanych uległa naciskowi. Czy ubocznym i niezamierzonym skutkiem badania nie było pogłębienie się u osób i tak już uległych poczucia zależności od innych, kiedy już uświadomiono im po zakończeniu eksperymentu fakt ulegania? Albo czy badani w eksperymentach Milgrama[130] nie zmienili *in minus* swej samooceny, gdy po zakończeniu badania objaśniono im faktyczny sens sytuacji i ich czynności?

Większość eksperymentów dokonywanych na ludziach i małych grupach nie nasuwa takich pytań etycznych. Bodźce są dla badanych niezbyt ważne, w każ-

[128] W. Dennis, Infant development under conditions of restricted practice and of minimum social stimulation, „Genetic Psychology Monographs" 1941, vol 23, nr 1.

[129] S. E. Asch, Nacisk grupy na modyfikacje i wypaczenia sądów, w: A. Malewski (red), Zagadnienia psychologii społecznej, Warszawa 1962.

[130] S. Milgram, Behavioral study of obedience, „Journal of Abnormal and Social Psychology" 1963, vol. 67, nr 4; tegoż: Some conditions of obedience and disobedience to authority, w: A.G. Miller (ed.), The social psychology of psychological research, New York 1972.

dym razie nie są dotkliwe bez ich dobrowolnej zgody, ich skutki są odwracalne i znikają wraz z zakończeniem eksperymentu. Grupy badane tworzy się *ad hoc* dla celów doświadczenia i rozwiązuje się po jego przeprowadzeniu. Jeśli eksperymentuje się na grupach naturalnych, to zakres manipulacji zawęża się jeszcze bardziej. Co więcej, udokumentowane odpowiedzi na pytania etyczne postawione pod adresem cytowanych wyżej eksperymentów nie zawsze usprawiedliwiają niepokój etyczny. Badania empiryczne nad możliwością pojawienia się negatywnych następstw manipulacji w eksperymencie Milgrama dla samooceny wskazują, że następstw takich, mimo obaw, brak. Niemniej jednak już samo pojawienie się tych wątpliwości sygnalizuje, że eksperymenty owe bliskie są punktom granicznym (a może wręcz są nimi), poza które manipulacja w naukach o zachowaniu człowieka wyjść nie może, jeśli nie chce narażać się na otwarty konflikt z panującymi systemami wartości ludzkich.

Wreszcie, samo eksperymentowanie na ludziach, bez względu na jego faktyczną czy tylko domniemaną, doraźną czy trwałą szkodliwość manipulacji, nie jest wolne od pewnych dwuznaczności etycznych.

Do tej kategorii zaliczyć wypada opory przeciw eksperymentowaniu z ludźmi wywodzące się z postawy zwanej potocznie kompleksem królika doświadczalnego, w literaturze anglosaskiej zwanej kompleksem świnki morskiej — oba te gatunki to standardowe zwierzęta doświadczalne. Margaret Mead tak charakteryzuje tę postawę uwzględniając wyniki badań nad wyobrażeniami społecznymi o nauce i uczonych w Ameryce: Nazywając człowieka poddanego eksperymentowi świnką morską, stosunek między eksperymentatorem a badanym „przyrównuje się do stosunku władzy, w którym istoty ludzkie sprowadzono do poziomu zwierząt eksperymentalnych. W zwierzętach eksperymentalnych widzi się zwierzęta wyhodowane dla celów* eksperymentu, pozbawione swobody, męczone i zabijane. Ich los pokazuje w ten sposób straszliwe granice, ku którym eksperyment pójść może. Istoty ludzkie umieszczone w sytuacji zwierząt laboratoryjnych doprowadza się do stanu bezsilności, pozbawia indywidualnych praw i godności”[131].

Wydaje się jednak, że argument posługujący się tą analogią ma charakter tradycyjny i nieoświecony — posługuje się perswazyjnymi skojarzeniami a nie analizuje rzeczywistej praktyki eksperymentalnej. Pojawia się on zapewne częściej w kręgach osób o słabszym kontakcie z wykształceniem i nauką a jego siła hamująca eksperymentowanie jest słaba — powstrzymywać może on co najwyżej przed uczestnictwem w eksperymentach w roli badanych a nie przed podejmowaniem samych eksperymentów przez badaczy.

O wiele poważniejsza i bardziej wyrafinowana jest argumentacja etyczna przeciw eksperymentom z ludźmi podnosząca negatywne strony ukrywania przed badanym faktu lub celu eksperymentu; zabieg ten jest standardową metodą aranżacji i prowadzenia badań eksperymentalnych. Argumentację tę rozpatrzymy dokładniej, gdyż w odróżnieniu od poprzednich, milcząco uznawanych bądź milcząco odrzucanych, jest ona kontrowersyjna i dyskutowana.

Zrozumiałym samo przez się warunkiem sensowności eksperymentu jest bliskość sytuacji eksperymentalnej do sytuacji pozaeksperymentalnych, na które będzie się ekstrapolować wyniki badania. Z tego względu najlepsze byłyby eksperymenty *in vivo* — w sytuacjach naturalnych. Badani nie zdają sobie wówczas sprawy ani z tego, że sytuacja w której funkcjonują jest traktowana jako eksperymentalna ani z tego, że bodźce, na które reagują mają charakter poznawczy. Trudność polega jednak na tym, że o sytuacje takie nie jest łatwo, wachlarz manipulacji dopuszczalnych skromny, możliwości kontroli małe. Toteż powszechnie eksperymentuje się *in vitro* — w sztucznych sytuacjach i na sztucznych grupach laboratoryjnych.

W obu rodzajach eksperymentowania poznawczy charakter sytuacji lub tylko rzeczywisty cel doświadczenia powinien być ukryty przed badanymi, sytuacje sztuczne powinni traktować jako naturalne. Również w obydwu przypadkach za-

[131] M. Mead, Research with human subjects: a model derived from anthropological field practice, w: P. A. Freund (ed.), Experimentation with human subjects, London 1972.

miast tworzyć sytuacje rzeczywiste, o których mówią testowane hipotezy, przekonuje się badanych, że w takich sytuacjach się znajdują. Manipulację przedmiotami — elementami sytuacji zastępuje więc manipulacja symbolami, głównie językowymi. Umiejętne ukrycie przed badanymi faktu eksperymentu lub jego rzeczywistego celu powinno przy tym mieć skutki równe wyprodukowaniu sytuacji rzeczywistych — wedle zasady współczynnika humanistycznego bowiem zachowania ludzkie wyznaczone są nie przez obiektywne sytuacje, ale przez ich subiektywne definicje, w konsekwencji „…ludzie kierują się raczej błędami wyznawanymi, a nie prawdą, której nie znają”[132], Co więcej, mistyfikacja tego rodzaju umożliwia uzyskanie poważnych korzyści praktycznych, metodologicznych i etycznych. Tak więc:

Mistyfikacja rozszerza repertuar eksperymentów możliwych, gdyż często ujawnienie przed badanymi eksperymentalnego charakteru sytuacji po prostu uniemożliwiałoby przeprowadzenie doświadczenia. Eksperyment Ascha może być tutaj przykładem. Innym razem pozwala ona zmniejszyć, nieraz do bardzo małych rozmiarów, koszty eksperymentowania. Często bowiem tworzenie sytuacji rzeczywistych byłoby zbyt kosztowne jak na moralną cenę samej tylko mistyfikacji albo byłoby złem nieporównanie większym niż tylko wprowadzenie w błąd. Tu za przykład niech znów posłuży eksperyment Milgrama: nie widać najmniejszego, poza dążeniem do prawdomówności, powodu, dla którego badacz miałby umożliwić badanym aplikowanie innym faktycznie bolesnych wstrząsów elektrycznych.

Wreszcie, mistyfikacja przyczynia się do zapewnienia badanym poczucia realizmu sytuacji, co utrudnia im dostarczanie danych pseudopotwierdzających lub pseudoobalających sprawdzane hipotezy. Posiadanie przez badanych świadomości uczestnictwa w eksperymencie, znajomość jego celów a szczególnie testowanych hipotez prowadzi w pewnych środowiskach do zachowań umyślnie niezgodnych, a częściej do umyślnie zgodnych z domniemywanymi hipotezami; wartość danych z takich badań jest mała. Zjawisko to każe ukrywać przed badanymi prawdziwy cel eksperymentu. Samo przemilczenie prawdy jednak niewiele daje, gdyż badani imputować mogą eksperymentatorowi cele własne, co gorsza jemu nie znane i czynić je elementami definiującymi sytuacje i wyznaczającymi reakcje. Toteż podaje się im fałszywy cel doświadczenia — przez to chce się ich zmylić a zniekształceniom nadać kierunek kontrolowany. Dokładniej o zjawiskach tych będzie mowa dalej.

Mistyfikacja eksperymentalna, jakkolwiek usprawiedliwiona — jak widać — poważnymi racjami naukowymi, budzi jednak opory motywowane innymi racjami kulturowymi. Najogólniej rzecz ujmując, ukrywanie prawdy o eksperymencie ma służyć ujawnieniu prawdy o świecie, uzyskaniu wiedzy pozwalającej przewidywać zachowania ludzkie w kontekstach pozaeksperymentalnych. Ale prawda tak pojęta nie tylko nie jest jedyną czy najważniejszą wartością kultury, ale i nie jest nawet jedyną odmianą prawdy jako wartości. Ta bowiem jako prawdomówność jest także fundamentem życia zbiorowego: jako wartość instrumentalna umożliwia niezbędne do kooperacji zaufanie, jako wartość samocelowa jest składnikiem ogólnej zasady szacunku dla osobowości. Prawdomówność w sensie odpowiedzialności za słowo zajmuje też szczególne miejsce w kulturze nauki — jest ona warunkiem koniecznym realizacji jej celu podstawowego tj. formułowania zdań prawdziwych czyli zgodnych z faktami[133].

Powstaje więc możliwość konfliktu między dążeniem do prawdy naukowej a postulatem prawdomówności. A stąd eksperymenty opierające się dla dobra nauki na kłamstwie artykułowanym bądź tylko kłamstwie przez przemilczenie budzą opory motywowane nie tylko szkodliwością manipulacji niemożliwych bez mistyfikacji.

Wskazuje się więc na negatywne skutki mistyfikacji eksperymentalnej dla badanych. Samo wprowadzenie w błąd ma naruszać istotne prawo jednostki do nie-

[132] B. Malinowski, Szkice z teorii kultury, Warszawa 1958, s. 149.

[133] O prawdzie i prawdomówności jako wartościach kultury traktuje artykuł pod tym tytułem w „Studiach Filozoficznych” 1966, nr 2.

podlegania manipulacji, ma sprowadzać ją do roli przedmiotu, niszczyć jej zaufanie do ludzi i instytucji. O jednym z eksperymentów Rokeacha dokonanym bez zgody i wiedzy badanych pisał z oburzeniem S. E. Miller, iż „stanowi on zamach na prawa człowieka a mianowicie prawo do niepodlegania manipulacji i eksperymentowaniu bez wiedzy i zgody. Sądzę — kontynuował — że jednym z najbardziej zasadniczych aspektów rozwiniętej kultury jest to, że obywatel może zakładać, iż w zwykłej, codziennej działalności będzie traktowany szczerze i godnie, i że w ogóle może zaufać ludziom, z którymi ma do czynienia”[134]

Dalej, dowodzi się istnienia ujemnych wpływów eksperymentowania z ukryciem celu na osobowość badaczy. Margaret Mead np. utrzymuje, że u badaczy je praktykujących powstaje przyzwyczajenie manipulowania i wprowadzania w błąd, zanika wrażliwość i samowiedza moralna, ustępując miejsca postawom cynicznym oraz przekonaniom o wadze i posłannictwie własnej roli i badań — mniemania te pozwalają Usprawiedliwiać osądzaną procedurę[135]. Podobne zjawisko zasygnalizował także Kelman. Zaniepokoiło go bezrefleksyjne stosowanie mistyfikacji eksperymentalnej. „Zbyt często wprowadzania w błąd używa się nie jako środka ostatecznego, lecz jako rzeczy oczywistej. Wydaje się — pisał on — że postawa nasza jest taka oto: jeśli możesz oszukać, po co mówić prawdę”[136].

Następnie, masowe eksperymentowanie z ukryciem celu, ujawnianym po zakończeniu doświadczenia ma negatywne skutki dla statusu badaczy jako kategorii zawodowej. Zauważono, że obniża ono prestiż psychologów i psuje ich reputację, tworzy stereotyp psychologa który kłamie, podrywa zaufanie do tego zawodu[137]. I to również, że eksperymentowanie tego rodzaju obniża zaufanie do laboratorium i wartość języka w społeczności naukowej[138]

W końcu, widzi się ujemne wpływy mistyfikacji eksperymentalnej dla szerszej niż naukowa kultury. Instytucjonalizacja i uświęcenie kłamstwa w badaniach naukowych ma być wkładem nauki w procesy przemiany człowieka w obiekt manipulowalny i manipulowany, jakie mają się dokonywać we współczesnych społeczeństwach masowych, przy czym myśli się tu o Ameryce[139].

Powstają pytania o słuszność i. skuteczność tych argumentów. Rozważymy je w porządku ekspozycji.

Jeśli idzie o ujemne skutki mistyfikacji dla badanych, to zauważyć trzeba, że zabiegiem równie standardowym, jak ona, jest demistyfikacja po zakończeniu eksperymentu. Ujawnienie faktu i celu eksperymentu oraz faktycznej roli badanych traktowane jest jako obowiązek etyczny, choć dostarcza również materiału naukowego. Ukazanie faktycznego charakteru sytuacji oraz szkód poniesionych lub wyrządzonych przez badanych ma nie tylko naprawić „zło”, ale też dać pewne wartości dodatnie — jest więc ono aktem rekompensaty moralnej. „Słuszną zasadą, której — zdaniem Kelmana — należy przestrzegać jest to, że badany nie powinien opuszczać laboratorium z większym lękiem lub niższą samooceną niż ta, z którą wszedł do laboratorium. Co więcej, winien on być w pewien sposób wzbogacony przez doświadczenie — powinien odchodzić z przekonaniem, że czegoś się nauczył, coś zrozumiał czy jakoś się rozwinął”[140]. Jakkolwiek jest możliwe, że sama demistyfikacja produkuje skutki ujemne, bo ujawniając badanym negatywne strony ich osobowości, takie jak konformizm, naiwność, sadyzm, ślepe posłuszeństwo, obniża im samoocenę, to jednak badania podjęte specjalnie w celu sprawdzenia tych domysłów nie wykazały ani doraźnych, ani trwałych trauma-

[134] S. E. Miller, Psychology experiments without subjects consent, „Science” 1955, vol. 152, nr 3718, s. 15,

[135] Mead, op. cit.

[136] H. C. Kelman, Time to speak: on human values and social research, San Francisco 1968.

[137] Zob. M. T. Orne, On the social psychology of the psychological experiment, w: Miller (ed.), op. cit. s. 237.

[138] Zob. D. T. Campbell, Prospective: artifact and control, w: R. Rosenthal, R.L. Rosnow (eds.), Artifact in behavioral research, New York 1969.

[139] Zob. Kelman, op. cit.

[140] H. C. Kelman, Deception in social research, w: I. Horowitz, (ed.), The uses and misuses of social sciences, New York 1971.

tycznych skutków demistyfikacji poeksperymentalnej dla osobowości badanych. Przeciwnie, np. w wypadku doświadczenia Milgrama badani w ogromnej większości wyrazili zadowolenie, że mogli wziąć udział w niecodziennej a pomysłowej sytuacji. Co więcej odsetek osób bardzo z tego zadowolonych był wyraźnie większy wśród posłusznych niż wśród opornych![141] Czy mistyfikacja eksperymentalna sprowadza badanego do manipulowanej rzeczy? Wątpliwe, jeśli zauważyć, że wiedza uzyskana przezeń w eksperymencie zwiększa jego możliwości wyborów niezależnych w przyszłych sytuacjach życiowych.

Jeśli idzie o charakter wpływów eksperymentowania z użyciem mistyfikacji na świadomość moralną badaczy, to niestety nie znamy ani nie słyszeliśmy o badaniach empirycznych nad tym fenomenem a tylko dane systematyczne pozwoliłyby ocenić krytycznie trafność argumentów Margaret Mead. Póki co, zdrowy rozsądek nakazuje być wobec nich sceptycznym — wydają się one przesadzone.

Poważniejszy charakter mają tezy o ujemnych skutkach mistyfikacji eksperymentalnej dla roli społecznej badaczy. Ale i tu zanotować warto, że w życiu społecznym istnieją liczne sytuacje, w których nakaz prawdomówności bezwzględnej nie obowiązuje a nawet istnieje etyczny zakaz mówienia prawdy. Sąd, szpital, wojna — są takimi miejscami, gdzie ukrywanie prawdy jest akceptowane dla obrony własnej, zmylenia nieprzyjaciela, oszczędzenia cierpień. Taką enklawą w kulturze zorientowanej na rozwój nauki może być i laboratorium. Można także zauważyć, że nie ma wcale zgody co do tego, że uczestnictwo w badaniach z ukryciem celu wzmaga podejrzliwość badanych wobec następnych eksperymentatorów.

Wreszcie, argumentowi o ujemnym wpływie mistyfikacji i manipulacji eksperymentalnej na kulturę, można przeciwstawić uwagę następującą: Istnieją dziedziny, jak np. psychoterapia, które posługują się tymi narzędziami i nie budzi to niepokojów moralnych, gdyż służą one celowi chwalebnemu tj. zwiększeniu wolności. Tak można spojrzeć i na eksperymentalne badania społeczne. Zanotowaliśmy już, że uczestnictwo w eksperymentach rozszerza swobodę wyborów u badanych. Natomiast wiedza uzyskana eksperymentalnie może być narzędziem obrony przed manipulacją dla wszystkich, którzy ją posiadają. Badania perswazji, konformizmu i posłuszeństwa mogą ułatwić utrzymanie niezależności sądów, nonkonformizm i nieposłuszeństwo.

Nie widać zatem dostatecznych racji etycznych, dla których badacze zachowań ludzkich mieliby powstrzymywać się przed eksperymentowaniem z mistyfikacją. Postawa niedogmatyczna w sprawach wartości połączona z rachunkiem strat i zysków nakazuje śmiało je podejmować.

Co innego jednak słuszność argumentów a co innego skuteczność. Wyżej przytoczone argumenty przeciw wprowadzaniu w błąd w eksperymencie nie są li tylko rojeniami pięknoduchów. Kompetentni obserwatorzy notują, że obiekcje te kształtują postawy znacznej części badaczy i tworzą faktyczną, choć nie tak sztywną jak naganność bodźców i ich następstw, barierę eksperymentowania na ludziach i grupach społecznych. „Gdybyśmy wyliczali moralne problemy psychologii, to tych, co robią eksperymenty wymagające wprowadzenia w błąd, umieścilibyśmy daleko niżej od tych, co powodowani skrupułami etycznymi, robią ich zbyt mało lub nie robią ich wcale — pisze McGuire o amerykańskich psychologach społecznych. Niepokojącym występkiem w zawodzie psychologa jest nie nadużywanie a niewykonywanie swoich obowiązków. Wydaje się nam, że gdyby nagle anioł śmierci spadł na naszych kolegów, to większość z nich zastałby w bezczynności a nie w grzechu”[142].

[141] Zob.: S. Milgram, Issues in the study of obedience, w: Miller (ed.), op. cit.; K. Ring, K. Wallston, M. Corey, Mode of debriefing as a factor affecting subjective reaction to a Mil- gram-type obedience experiment: an ethical inquiry, w: J. Katz (ed.), Experimentation with human beings, New York 1972.

[142] W. J. McGuire, Suspiciousness of experimenter’s intent, w: R. Rosenthal, R.L. Rosnow (eds.), op. cit., s. 53.

9.4. REGULACJA KONFLIKTÓW WARTOŚCI

Eksperymentowanie w naukach społecznych, jak i w naukach o człowieku w ogólności, rodzi pewne pytania etyczne, takie jak: Jak pogodzić interesy nauki z autonomią dobra i woli jednostek? Jaka jest granica ryzyka eksperymentalnego? Pod jakimi warunkami wolno je podejmować? Etyka nauki usiłuje wypracować zasady będące odpowiedziami na te pytania. Żąda się od nich, by były jednoznaczne i zupełne oraz by zabezpieczały wartości ludzkie, nie hamując zarazem rozwoju nauki. Zasady te ujmują w kanon pewne pospolite intuicje moralne, którym dała swego czasu wyraz w esencjonalnej wypowiedzi Antonina Kłoskowska. „Eksperyment dokonany na człowieku — pisała ona — i mogący przynieść szkodę jego psychicznym lub fizycznym władzom usprawiedliwić można jedynie pod warunkiem, że przedmiot eksperymentu zdaje sobie sprawę z grożącego niebezpieczeństwa i dobrowolnie je podejmuje. Badaczowi, który operuje tak swoistym materiałem jak człowiek nie wolno zapominać o nieprzekraczalnej granicy określającej swobodę jego manipulacji. Granicę tę stanowi dobro społeczeństwa i jednostki”[143].

Budowa kodeksów eksperymentowania uściślających takie intuicje jest zadaniem niezwykle trudnym, jeśli mają one w sposób optymalny spełniać stawiane im zadania i warunki.

Po pierwsze, brak jest jednolitych „zasad korespondencji” łączących terminy, które w kodeksach tych miałyby występować z konkretnymi sytuacjami, do których oceny chciałoby się te kodeksy stosować. Cóż to bowiem znaczy „zgoda dobrowolna” w odniesieniu do praktyki eksperymentalnej? Odpowiedź, że można mówić o niej wtedy, gdy decydent świadom jest konsekwencji swego wyboru i podejmuje ryzyko bez czyjegokolwiek przymusu, dopiero daje początek wieloznacznościom Badany nie zawsze może zdawać sobie sprawę z następstw swego wyboru, bo bywa, że i sam eksperymentator nie potrafi ich przewidzieć; przebieg i ryzyko doświadczenia mają charakter zbyt specjalistyczny by badany nie dysponujący wiedzą fachową pierwsze mógł zrozumieć a drugie ocenić, poza tym w wielu eksperymentach zakłada się niewiedzę badanych o celach lub fakcie eksperymentu. Również „przymus” należy do terminów klasycznie rozciągliwych. To nie tylko przemoc fizyczna, ale i konieczność praktyczna, presja obiecanej nagrody czy darowanej kary. Ponadto, kto miałby kwalifikować sytuację jako przymusową: doświadczający nacisku czy ktoś inny przymierzający jego położenie do **pew**nego ideału etycznego? W dodatku, nawet w pełni świadoma i dobrowolna zgoda nie upoważnia do stosowania dowolnie ryzykownych manipulacji eksperymentalnych. Nie mniejsze trudności napotyka też operacjonalizacja pojęć dobra jednostki i dobra społeczeństwa: kto ma orzekać, co jest dobre dla jednostki (ona sama czy etyk?) i co jest dobre dla społeczeństwa?

Po drugie, zaangażowane w eksperyment wartości są niewspółmierne. Wartość wiedzy z jednej strony a dobro jednostek z drugiej są to jakości zakresowo nie tożsame: w przypadku konfliktu nie sposób ułożyć alternatyw na jednej skali i odczytać wyniku wyznaczającego wybór. Nie sposób też ustalić powszechnej preferencji dla którejś z wartości, gdyż wiodłoby to albo do zahamowania rozwoju wielu nauk albo do zbrodni w imię tego rozwoju — w istocie same owe preferencje zależą od indywidualnych cech sytuacji. Niewspółmierność ta polega i na tym, że kto inny ponosi koszty eksperymentu a kto inny czerpie zeń korzyści. Doraźne dolegliwości i możliwe szkody ponoszą badani — korzyści z rozwoju nauki osiąga całe społeczeństwo. Interesy obu podmiotów nie są tożsame, z tej racji kodeksy eksperymentowania muszą odwoływać się do jakiejś koncepcji wzajemnego stosunku jednostki i społeczeństwa.

Nie mniej istotne trudności nasuwa realizacja postulatów formalnych wysuwanych pod adresem kodeksów eksperymentowania. O jednoznaczności była już mowa. Podobnie jest z zupełnością: ponieważ kodeksy nie mogą raz na zawsze określić preferencji dla wartości konkurujących, to muszą one podawać wartość pre-

[143] A. Kłoskowska, Droga do społeczeństwa, „Wiedza i Życie” 1949, nr 11.

ferowaną i granice ustępstw dla bardzo wielu typów sytuacji. W tych warunkach mogą one zestandaryzować tylko sytuacje typowe, omijając pozostałe, a szczególnie te, które życie i postęp nauki ze sobą niosą.

Teoretyczna niemożliwość rozstrzygnięcia z góry i raz na zawsze dyskutowanego rodzaju konfliktu między interesami nauki a wartościami ludzkimi sprawia, że jest on ciągle żywy i nadaje świadomości moralnej uczonych rolę jego regulatora[144].

Rozważania tego rozdziału mogły sprawić wrażenie, ze zakres zjawisk społecznych badanych eksperymentalnie jest bardzo wąski. Nie było to celem autora. Zakres ten jest bowiem bardzo rozległy przedmiotowo, ale zarazem bardzo wąski w stosunku do dających się pomyśleć możliwości technicznych.

10. TZW. SZTUCZNOŚĆ SYTUACJI EKSPERYMENTALNYCH JAKO BARIERA EKSPERYMENTU

Aranżacja sytuacji eksperymentalnych i kontrola działania bodźców nie tylko prowadzi do poznania rzeczywistości lecz także odkształca ją. W konsekwencji,' barierą eksperymentowania staje się — obok niemożliwości tworzenia dowolnych sytuacji i wprowadzania dowolnych zmian — również ograniczona wartość uogólniania zależności ustalonych w sytuacjach zaaranżowanych na warunki pozaeksperymentalne, „naturalne''. Poprzednie bariery w wielu dających się pomyśleć przypadkach eksperyment uniemożliwiają — bariera o której mowa ogranicza natomiast jego wartość.

Eksperyment winien być zaprojektowany i przeprowadzony tak, by eksperymentator mógł trafnie skojarzyć bodziec z reakcjami, by — innymi słowy — nie przypisywał zaobserwowanych zmian swym manipulacjom, gdy faktycznie wywołały je nieplanowane czynniki uboczne. Postulat ten to postulat kontroli. Nakazuje on izolację układu badanego od wpływów zewnętrznych albo przynajmniej analityczne rozdzielenie wpływów zmiennych eksperymentalnych i zmiennych ubocznych. Jednakże sytuacja eksperymentalna nie interesuje badacza dla niej samej — interesuje go ona jako pewien przypadek szerszej klasy zjawisk, na którą chce on przenieść wyniki eksperymentu Stąd również dla dobra eksperymentu powinna być podobna do sytuacji pozaeksperymentalnych tak dalece, jak tylko jest to możliwe. Te dwa postulaty są — jak widać — sprzeczne. Co jest dobre dla trafności wewnętrznej, nie musi być dobre dla trafności zewnętrznej a ściśle dla tego jej wymiaru, który nazywa się trafnością ekologiczną lub sytuacyjną i *vice versa.*

Zagadnienie trafności ekologicznej wyników eksperymentu — czyli możliwości uogólniania tych wyników na konteksty pozaeksperymentalne — wygodnie jest rozważać wówczas, gdy sformułuje się je w terminach różnic zachowań ludzkich w warunkach eksperymentalnych i nieeksperymentalnych. Postawmy więc pytanie: dlaczego to zachowania ludzi w kontrolowanych sytuacjach eksperymentalnych bywają różne od zachowań w warunkach nieeksperymentalnych? Odpowiedzi na to pytanie szukać trzeba w specyficznych właściwościach grup biorących udział w eksperymencie, w osobliwościach psychospołecznej sytuacji badania eksperymentalnego oraz w samej naturze metody eksperymentalnej.

[144] Niezwykle bogatą dokumentację dyskusji nad etyczną stroną eksperymentowania na ludziach przynosi 1100-stronicowy wybór J. Katza, Experimentation with human beings, op. cit. Ważnym głosem w tej debacie jest artykuł B. Barbera, Eksperymentowanie na ludziach, „Zagadnienia Naukoznawstwa" 1970, nr 1. Próbę sformułowania zasad etycznych eksperymentów społecznych zawiera książka A. M. Rivlin, P. M. Timpane (ed-s.), Ethical and legal issues of social experimentation, Washington 1975.

10.1. SPECYFIKA GRUP EKSPERYMENTALNYCH

Specjalne własności grup, na których prowadzi się badania eksperymentalne stanowią pierwsze źródło odmienności zachowań eksperymentalnych względem zachowań nieeksperymentalnych. Nie musi myśleć o nich ten badacz, który przedmiotem manipulacji czyni grupy naturalnie istniejące i stosuje tylko takie bodźce, jakie mogłyby się w życiu tych grup zdarzyć. W eksperymentach takich badanie wmontowane jest w tryb normalnego funkcjonowania grupy, a jej członkowie nie muszą sobie nawet zdawać sprawy z tego, że są przedmiotem badania. Warunki te mogą być spełnione w eksperymentach terenowych a zwłaszcza w eksperymentach społecznych. Większa część wyników eksperymentalnych uzyskanych przez socjologię małych grup i psychologię społeczną pochodzi jednak z eksperymentów laboratoryjnych, dokonywanych na grupach tworzonych specjalnie na potrzeby badania. Grupy laboratoryjne różnią się od grup naturalnych wieloma istotnymi dla zachowania własnościami.

1. Grupy laboratoryjne utworzono doraźnie na użytek badania. Nie istniały one przed jego podjęciem, nie posiadają więc — inaczej niż grupy naturalne — ani trwałego składu, ani trwałej, struktury. Są to raczej grupy tylko z nazwy — bardziej agregaty niż grupy w sensie socjologicznym. W konsekwencji, członkowie ich nie będą w swych zachowaniach eksperymentalnych brać pod uwagę wielu podstawowych zmiennych charakterystycznych dla życia grup naturalnych, jak np. statusu partnerów, ustalonych kanałów komunikacji, nade wszystko zaś ustalonych wzorów interakcji społecznych[145].

2. Małe grupy naturalne pełnią szereg ważnych funkcji wobec jednostek i zaspokajają wiele ich ważnych potrzeb psychicznych. Kształtują osobowość, uczucia i ideały, wpajają postawy społeczne, zaspokajają potrzeby afiliacji, bezpieczeństwa i aktywności społecznej. Tymczasem „chcąc sformułować definicję małej grupy na podstawie badań, których była przedmiotem, trzeba by poprzestać na stwierdzeniu, że jest to jakakolwiek zbiorowość, w której istnieje możliwość kontaktów” — pisze Łoś-Bobińska[146]

3. Grupy laboratoryjne znajdują się w niecodziennym otoczeniu — laboratorium psychologiczne jest miejscem poniekąd wyjątkowym. Pozbawione są dla dobra kontroli możliwości interakcji z otoczeniem poza laboratoryjnym. Zadania, jakie się im stawia, nie są podobne do zadań spotykanych na co dzień.

4. Wprawdzie zasada realizmu eksperymentalnego wymaga by badani postrzegali stworzone przez eksperymentatora sytuacje jako „rzeczywiste” i reagowali na nie tak jak w warunkach codziennych, to jednak sama owa zasada jest nierealistyczna. Badani zwykle zdają sobie sprawę z eksperymentalnego, quasi-realnego charakteru sytuacji, w jakiej się znajdują, choć rzadziej znają faktyczny cel doświadczenia czy swoją rolę w badaniu, ukrywane przed nimi z ważnych powodów metodologicznych. Wspomniane w poprzednim punkcie czynniki tj. otoczenie, izolacja i charakter zadań znakomicie uprzytomniają badanym ich faktyczną sytuację tj. uczestnictwo w eksperymencie, znajdowanie się w sytuacji, w której obowiązują odmienne reguły postępowania. O następstwach świadomości uczestniczenia w eksperymencie będzie mowa dalej.

W konsekwencji pojawia się pytanie, czy bodźce stosowane przez eksperymentatora w warunkach laboratoryjnych są w stanie pobudzać te same motywacje, które kierują zachowaniami ludzkimi w naturalnych sytuacjach życiowych. Nie ulega wątpliwości, że waga następstw decyzji, koszty pomyłek itp. są niebywale różne w obu rodzajach warunków. Nie kwestionuje się też, że badani nie biorą pod uwagę odleglejszych skutków swoich działań — grupy kończą wszak życie wraz z zakończeniem badania. Wreszcie istnieje zgoda co do tego, że motywa-

[145] Zob. Th. E. Drabek, J. E. Haas, Realism in laboratory simulation: myth or method?, „Social Forces” 1967, vol. 45, nr 3

[146] M. Łoś-Bobińska. Pojęcie grupy społecznej w teoriach i badaniach małych grup, w: Moralność i społeczeństwo, Warszawa 1968, s. 166.

cje działań laboratoryjnych nie są na ogół tak intensywne, jak motywacje działań życiowych. Na tym jednak zgoda się kończy.

Dla jednych różnice między bodźcami eksperymentalnymi a okolicznościami życiowymi oraz między pobudzanymi przez jedne i drugie motywacjami mają charakter wyłącznie ilościowy — są różnicami stopnia intensywności i komplikacji bodźców i motywacji. Dla drugich pogląd ten jest zasadniczo nie do przyjęcia. Ich zdaniem, poziomy organizacji czynności i stopień uczestnictwa różnych mechanizmów regulujących zachowanie w obu rodzajach sytuacji różnią się od siebie zasadniczo.

Różnic między zarysowanym stanowiskiem nie da się sprowadzić do sposobu wyrażania się. Konsekwencje z nich wypływające są poważne. Pierwsze stanowisko prowadzi do przekonania, że ludźmi w laboratorium i w życiu codziennym kierują te same, co najwyżej modyfikowane, prawa. Drugie stanowisko wiedzie do poglądu, że sytuacje życiowe i sytuacje eksperymentalne są podstawieniami poprzedników różnych praw. Pogląd taki wymagałby od jego zwolenników stanowczego opierania się pokusie stosowania praw potwierdzonych w laboratorium do wyjaśnienia i przewidywania zachowania w sytuacjach życiowych. Wątpliwe, czy akceptują oni tę konsekwencję w swej praktyce badawczej.

10.2. EKSPERYMENT JAKO SYTUACJA SPOŁECZNA

Eksperyment nie odbywa się w próżni społecznej. Przeprowadza się go w pewnym szerszym kontekście społeczno-kulturowym a i sama sytuacja badania jest pewną sytuacją społeczną. Z tych banalnych konstatacji wynikają niezmiernie ważne konsekwencje dotyczące zachowania się badanych, roli eksperymentatora a w efekcie — i samych wyników.

Badani w eksperymentach socjo- i psychologicznych nie są biernymi obiektami manipulacji eksperymentatora, reagującymi wyłącznie na precyzyjnie przezeń wyizolowane bodźce, i to w myśl prostego schematu „bodziec — reakcja". Uczestniczą oni natomiast w całościowej sytuacji społecznej a bodźce są tylko pewnymi jej elementami. Co więcej, oddziałują one na ich zachowanie o tyle tylko, o ile są składnikami subiektywnych definicji sytuacji, wypracowanych przez badanych — owe definicje określają ponadto sens bodźców. Zachowania badanych są współwyznaczone przez wartości ich kultury i środowiska, a także przez sposób pojmowania roli badanego, potrzeby osobowościowe, cele imputowane eksperymentatorowi.

Eksperymentator nie jest jedynym twórcą sytuacji eksperymentalnej, a co więcej, sam jest jej częścią. Oddziałuje on na zachowania badanych nie tylko zaplanowanymi bodźcami, ale i całością zaangażowanych w eksperyment cech swej pozycji społecznej i osobowości. Jest on przedmiotem postaw badanych. Między nimi a badanymi zawiązują się elementarne stosunki społeczne.

W efekcie wyniki eksperymentu zależą nie tylko od zastosowanych bodźców, lecz także od wymienionych wyżej czynników, w szczególności od interakcji społecznych tworzących się w procesie badania.

Sformułowaniom tym nadają bardziej precyzyjny i sprawdzalny charakter wyniki badań z zakresu metodologii empirycznej zarówno dotyczące bezpośrednio eksperymentu, jak też innych metod opartych na komunikacji badacza z badanym. Ustalenia te każą przyglądać się dokładnie wynikom wielu eksperymentów z zakresu nauk o zachowaniu społecznym. ,

Sens sytuacji eksperymentalnej inny bywa dla badacza a inny dla badanego, jeśli ten ostatni zdaje sobie sprawę z tego, że ma ona charakter eksperymentalny właśnie. Jeśli nawet obaj zainteresowani są w wynikach eksperymentu, to różnie mogą je pojmować, co dla uzyskanego wyniku nie jest bez znaczenia. Wprawdzie obaj dążą do dostarczenia lub uzyskania danych dobrych tj. rzetelnych i zwykle

nie dokonują zafałszowań świadomych, to jednak dane te obarczone są błędami systematycznymi.

Powiedzieliśmy już, że zachowania badanych zależą nie tylko od bodźców, lecz także od wartości kultury, do której badani należą lub aspirują. Wzgląd na te wartości wyznacza pierwszy kierunek tendencyjności zachowań badanych: reakcje badanych są współwyznaczone przez ich sądy, o tym, jakie reakcje są społecznie aprobowane a jakie aprobowane nie są. I tak, wiadomo, że w społecznościach tradycyjnych, w których małe zróżnicowanie wartości jest normą kulturową a idea badań społecznych nie jest znana, w badaniach uzyskuje się jedynie odpowiedzi stereotypowe, konformistyczne wobec standardów kultury — odpowiedzi zindywidualizowane byłyby dewiacjami[147]. Wiadomo też, że w społeczeństwach podlegających szybkim zmianom, w których istnieje rozbieżność wartości propagowanych i uznawanych a zarazem, tak jak wyżej, układ ról „badacz — badany" nie stanowi części kultury, zaznacza się skłonność do deklarowania badaczowi wartości i ujawnienia zachowań oficjalnych a ukrywania prywatnych[148].

Dobrą ilustracją pierwszej zależności jest próba testu polegającego na jak najszybszym rozwiązaniu pewnego zadania, przeprowadzona przez Ascha wśród Indian z plemienia Hopi[149], którym ich apollińska kultura zakazuje podejmowania wzajemnego współzawodnictwa. Choć jedni rozwiązali zadanie szybciej a drudzy wolniej, to porozumiawszy się, rozwiązanie oznajmili wszyscy równocześnie. Przykładem drugiej, lepiej znanej tendencji może być choćby zjawisko zawyżania poziomu wykształcenia w badaniach ankietowych, spisach powszechnych i badaniach piśmienności, obserwowane tak w Polsce, jak i w innych krajach — deklaracje te są wyrazem aspiracji badanych do wartości publicznych[150].

Zarówno podane wyżej twierdzenia, jak i przykłady wzięte są z badań nieeksperymentalnych. Nie jest to jednak przeszkodą do odnoszenia ich do badań eksperymentalnych, albowiem często reakcje badanych w eksperymencie mają charakter zachowań werbalnych, a co więcej, pomiaru skutków działania bodźców dokonuje się często przy pomocy interrogacji. Zresztą standaryzujący wpływ wzorów kulturowych na reakcje eksperymentalne o charakterze społecznym trudno ilustrować danymi dotyczącymi wprost eksperymentów, gdyż eksperyment jako technika dosyć wyrafinowana, stosowany był głównie w jednorodnym i rozwiniętym środowisku kulturowym, co uniemożliwia określenie przez kontrast znaczenia wpływu środowiska na zachowania eksperymentu.

Na brak materiałów nie można natomiast narzekać wtedy, gdy przystępuje się — co właśnie czynimy — do pokazania wpływu rozmaitych aspektów samej społecznej sytuacji badania eksperymentalnego na jego przebieg i uzyskane wyniki. Zagadnienia te są od kilkunastu lat przedmiotem intensywnych badań amerykańskich psychologów społecznych. Badacze ci usiłują odpowiedzieć na takie oto pytania: Kim są badani w eksperymentach, czy wyróżniają się oni pod jakimiś względami z całej populacji? Czym kierują się Judzie podejmując decyzję o dobrowolnym wzięciu udziału w eksperymencie? Jakie są cechy sytuacji bycia przedmiotem badania eksperymentalnego? Jakie motywacje powodują ludźmi w czasie badania? Od czego zależy kooperatywne, obojętne czy też negatywne nastawienie badanych wobec eksperymentatora? W jakim kierunku idą niezamierzone, pozastymulacyjne oddziaływania eksperymentatora na badanych? Jakie cechy biofizyczne, osobowościowe i społeczne eksperymentatora określają skuteczność tego wpływu? Jakimi kanałami wpływ ów się dokonuje? Jaka jest ogólność środowisko-

[147] Zob. P. Pradervand, Problemy sytuacji społecznej wywiadów w badaniach nad stosunkiem ludności do planowania rodziny w krajach słabo rozwiniętych, w: Z. Gostkowski, J. Lutyński (red.), Analizy i próby technik badawczych w socjologii, Wrocław 1970, t. 3.

[148] Zob. Z. Gostkowski, O założeniach i potrzebie empirycznych badań nad technikami i procedurami w socjologii, w: Z. Gostkowski, (red.), Analizy i próby technik badawczych w socjologii, Wrocław 1966. t. 1.

[149] Zob. O. Klineberg, Rasa a psychologia, w: L. C. Donn i in. Rasa a nauka, Warszawa 1961, s. 83.

[150] Zob.: A. Sztaudynger-Kaliszewicz, Badania wiarygodności zeznań w Spisie Powszechnym, w: Gostkowski (red.), op. cit.; G. Carceles-Breis, Dwie techniki pomiaru piśmienności w krajach słabo rozwiniętych, w: Gostkowski, Lutyński, (red.), op. cit.

wa procesów psychospołecznych zachodzących w badaniach eksperymentalnych? Zreferujemy odpowiedzi na te pytania udzielane[151]

Wedle czyjegoś powiedzenia współczesna psychologia społeczna jest nauką o zachowaniu amerykańskich studentów młodszych lat psychologii. Oprócz wysokiej oceny wkładu badaczy amerykańskich w rozwój tej dyscypliny i oprócz złośliwości w opinii tej obecne jest także pewne twierdzenie opisowe. Co więcej, jeśli odnieść je do psychologii amerykańskiej, to jest ono w znacznym stopniu prawdziwe.

Smart i Schultz niezależnie od siebie sprawdzili to twierdzenie na kilkunastu, pochodzących z lat sześćdziesiątych, rocznikach czołowych periodyków publikujących wyniki badań eksperymentalnych tj. „Journal of Abnormal and Social Psychology" („Journal of Personality and Social Psychology") i „Journal of Experimental Psychology"[152]. Wnioski z ich badań były zbieżne. „Czytając nasze czasopisma — konkluduje Schultz — odbiera się dobitne wrażenie, że jedynym rodzajem ludzi, jakimi interesują się psychologowie, są studenci!". Tak np. żadne z badań przedstawionych w „Journal of Experimental Psychology" nie było oparte na próbie z całej populacji, a w drugim z pism badania takie należały do niezwykłych wyjątków. Natomiast w lwiej części badań (70 - 86%) rolę badanych przyjmowali na siebie studenci; w połowie przypadków byli to studenci pierwszych lat psychologii.

Wyjaśnienie tego stanu rzeczy jest mało skomplikowane — studenci a zwłaszcza studenci psychologii są najłatwiej dostępnymi obiektami: są pod ręką eksperymentatora a i powodów do brania udziału w eksperymentach mają więcej niż kto inny.

Opisane zjawisko samo w sobie nie byłoby niczym nagannym, byleby tylko studenci byli reprezentatywni dla całej zbiorowości, do której chce się odnosić wyniki badań. Rzecz w tym, że nie są. Wiek, poziom inteligencji, korelaty pochodzenia społecznego -wyróżniają studentów spośród ogółu. Stąd Smart sugeruje, że wiele wyników badań nad uczeniem się, zwłaszcza werbalnym, może mieć małą ogólność środowiskową dlatego, że zostało ustalone w badaniach osób, które znajdowały się u szczytu swych możliwości intelektualnych.

Zagadnienie reprezentatywności osób biorących udział w eksperymentach ma swój ciąg dalszy. Oto bowiem eksperymentatorzy bardzo często korzystają z dobrej woli ochotników, którzy sami zgłaszają się do badań. Powstaje pytanie o to, kim są ci ochotnicy. Istnieją odpowiedzi na nie oparte na badaniach. Rosenthal i Rosnow dokonawszy przeglądu takich badań[153], stwierdzili, że ochotnicy nie są reprezentatywni nawet dla ogółu badanych. Są oni młodsi niż reszta, bardziej wykształceni i inteligentni, mają wyższy status zawodowy, są mniej autorytarni, wykazują silniejsze motywacje intelektualne i silniejszą potrzebę uznania. Dla każdej z tych cech można znaleźć zmienną, dla której jest ona istotna. Zobaczymy dalej, że dwie ostatnie z tych cech są ważne dla postaw badanych w eksperymencie, obciążających jego wyniki.

Motywy, dla których badani a zwłaszcza studenci dobrowolnie biorą udział w eksperymentach są różne. Może im chodzić o uzyskanie drobnej sumy pieniędzy, oswojenie się z laboratorium potrzebne do własnych studiów, wyświadczenie uprzejmości eksperymentatorom, poznanie siebie samego, przeżycie czegoś nowego. Nie ma zgody wśród badaczy co do najważniejszego z motywów. Orne uznał zań „nadzieję i oczekiwanie [badanych], że badanie, w którym biorą udział wniesie pewien istotny wkład do nauki a w ostateczności może i do ludzkiej po-

[151] Najważniejszymi publikacjami zawierającymi te odpowiedzi są: R. Rosenthal, R. L. Rosnow (eds.), Artifact in behavioral research, New Jork 1969; A. G. Miller (ed.), The social psychology of psychological research, New Jork 1972. Na zamieszczone w nich studia często będziemy się powoływać. Już po napisaniu tej książki, ukazała się monografia I. Silvermana, The human subject in the psychological laboratory, New York 1977 i poświęcony tym sprawom numer pisma „Et cetera" 1977, vol. 34, nr 3; zob. zwłaszcza artykuły Rosenthala oraz Rosnowa i Davisa.

[152] D. P. Schultz, The human subject in psychological research, w: Miller (ed.), op. cit., autor ten referuje też wyniki Smarta.

[153] R. Rosenthal, R.L. Rosnow, The volunteer subject, New York 1975.

myślności w ogóle”[154]. Natomiast Schultz sądzi, że „tak wzniosły i szlachetny cel można przypisać tylko pewnej mniejszości badanych studentów”. Jego zdaniem tym, co popycha ludzi do laboratorium jest ich ciekawość[155]. Motywy te, podobnie jak i szczególne cechy badanych, okażą się ważne dla zachowania się badanych w eksperymencie a zwłaszcza dla jego interpretacji.

Takie to nietypowe i kierujące się nietypowymi motywami osoby przekraczają próg laboratorium psychospołecznego. Rzadko są one na tyle naiwne, że objaśnienia i instrukcje eksperymentatora wezmą dosłownie. Częstokroć mają one już za sobą pewne osobiste doświadczenia w uczestnictwie w eksperymentach, często też mają pewne już wyrobione nastawienie wobec tego eksperymentu lub wobec eksperymentowania w ogóle. Wiedza, jaką nabyli w toku studiów lub informacje uzyskane od kolegów są także ważnym elementem ich wyposażenia. Może mieć ono istotny wpływ na ich przyszłe zachowanie w eksperymencie. Utrudni ono eksperymentatorowi przekonanie ich o realności stworzonych sytuacji a im ułatwi samodzielną interpretację sytuacji eksperymentalnej. Istnieją badania pokazujące, jak doświadczenie w byciu przedmiotem eksperymentu wpływa na zachowanie w następnych eksperymentach. Wyniki ich nie są spójne, więc nie będziemy ich referować. W każdym jednak razie, osoba wchodząca do laboratorium psychologa Społecznego to nie jest *tabula rasa,* na której można napisać wszystko, co się żywnie podoba.

Obecnie trzeba dokładniej opisać samą sytuację eksperymentu. W klasycznym już ujęciu Rieckena sytuacja eksperymentu widziana jako szczególny rodzaj interakcji społecznej odznacza się następującymi własnościami:[156]

1. „Najważniejszą cechą eksperymentu jako sytuacji społecznej jest to, że stanowi on zaproszenie dla jednej osoby, by zachowywała się wedle wskazówek drugiej”. Badany przyjmuje to zaproszenie, gdyż spodziewa się pewnych nagród — wymieniliśmy je omawiając motywy uczestnictwa w eksperymentach. „W pewnym sensie zatem, badany może mieć swój własny zespół prywatnych powodów lub celów dla wzięcia udziału w eksperymencie”.

2. „Warunki zaproszenia do zachowania nie są zwykle dobrze określone”. Eksperymentator podaje dokładnie tylko czas, miejsce, długość trwania doświadczenia itd., natomiast cel badania — bądź co bądź najważniejszy w tym wszystkim — podaje tylko ogólnikowo. Badany sam posiada jakieś niedokładne wyobrażenia o eksperymentowaniu w ogóle lub o eksperymentach pewnego rodzaju, o zainteresowaniach badaczy, o tym jak ci starają się zainteresowania swe zaspokoić, o granicach zachowania właściwego itd.

3. Szczególną jest natura zaangażowanych w eksperyment postaci. W większości wypadków eksperymentatorem jest profesor, a badanym — student. Stosunek istniejący między tymi rolami ma społecznie określony charakter.

Eksperymentator będący profesorem i psychologiem „jest wszechwładną postacią. Ma on dwa rodzaje władzy — ma władzę nad studentami, jako profesor, członek nadrzędnej, oceniającej ich grupy; jako psycholog ma on moc wglądu w badanego. W dodatku, można mieć wobec niego ambiwalentne odczucia. Z jednej strony, jest on godny zaufania w tym sensie, że nie będzie świadomie ryzykował życia badanego, atakował go ukradkiem lub bezpośrednio. Z drugiej strony, zna tę samą sztuczkę, którą znają wszyscy nauczyciele: odgadywanie prawdziwego charakteru ucznia, ocenę jego wartości i wystawianie mu cenzurki”. Jest to obraz przerysowany — eksperymentatora można oszukać, wartość ryzykowanych w tej grze nagród nie jest wielka, jednakże „dla badanego, który gra ostrożnie i stara się uzyskać od profesora dobrą ocenę dla siebie, najlepszą strategią może być pozorna uległość i pozorna chęć współpracy”.

[154] M. T. Orne, On the social psychology of the psychological experiment: with particular reference to demand characteristics, w: Miller (ed.), op. cit. s. 236.
[155] Schultz, op. cit., s. 37.
[156] H. W. Riecken, A program for research on experiments in social psychology, w: N. F. Washburne (ed.), Decisions, values and groups, Oxford 1962, vol. 1, s. 28 - 32.

4. Eksperyment jako sytuacja społeczna jest — przynajmniej z punktu widzenia eksperymentatora — „oddzielony czasowo i przestrzennie od życia codziennego". „Eksperymentator chce zwykle użyć badanego jako pewnego przypadku zachowania, użyć go akurat jeden raz a potem zapomnieć o nim wszystko z wyjątkiem dostarczonych przezeń danych. Widzi on w badanym mniej lub bardziej typową istotę ludzką a w jego zachowaniu mniej lub bardziej typowy przypadek tego, co istota ludzka uczyniłaby w warunkach eksperymentu".

Z punktu widzenia badanego sytuacja eksperymentu może wyglądać zgoła inaczej. Badany może nie zaakceptować traktowania go jako tylko egzemplarza gatunku i będzie dążył do tego, by eksperymentator dostrzegł w nim pewną indywidualność — będzie więc chciał zyskać dla siebie pewne uznanie. Może on też nie wierzyć, że stanowisko eksperymentatora wobec jego zachowania wyprane jest z emocji — przeciwnie, może być przekonany, że emocjonalna postawa eksperymentatora wobec niego znajdzie wyraz w przyszłych jego kontaktach z eksperymentatorem i innymi osobami mającymi dostęp do danych, w ferowanych przez nich ocenach i rekomendacjach dotyczących toku jego studiów i kariery. „W każdym wypadku, pytanie badanego brzmi: Co on sobie o mnie pomyśli?".

5. Wreszcie, społeczną sytuację eksperymentu cechuje „jednostronny rozdział informacji". Eksperymentator zapraszając badanego do udziału w eksperymencie ukrywa przed nim nie tylko prawdziwy lub dokładny cel doświadczenia, lecz także „odpowiedzi poprawne" a nawet kryteria oceny jego zachowania. Zapewnia natomiast badanego, że nie ma reakcji złych lub dobrych a są tylko autentyczne. Jednakże badany „podejrzewa, że różne odpowiedzi są poprawne i złe w tym sensie, że ukazują go one eksperymentatorowi lub nie ukazują w takim świetle, w jakim chce się on (badany) pojawić, że istnieją takie reakcje, które podnoszą i takie, które obniżają jego wartość jako osoby. Fakt, że eksperymentator kontroluje dostępną badanemu informację i że nigdy nie ujawnia zupełnie, co próbuje odkryć i jak będzie oceniał to co zaobserwuje, w dużym stopniu jest przyczyną tego, że eksperyment ma charakter gry lub rywalizacji. Prowadzi on do zespołu czynności inferencyjnych i interpretacyjnych podjętych przez badanego po to, by przeniknąć tajemnicę eksperymentatora i dojść do pewnego pojęcia o znaczeniu, jakie nieujawnione kategorie reakcji mają dla eksperymentatora".

Z przedstawionej charakterystyki sytuacji eksperymentu wynika, 'że zachowanie badanego składa się — najogólniej mówiąc — z dwu faz: rozwiązania pewnego problemu i uwzględnienia uzyskanego rozwiązania przy wyborze reakcji eksperymentalnej.

Rozwiązanie problemu to odkrycie prawdziwego charakteru eksperymentu a dokładniej — sprawdzanej hipotezy. Można zapytać, czym badany kieruje się w swych próbach rozwiązania tego zadania. Istnieje wiele wskazówek pozwalających mu odgadnąć cel eksperymentu. Obejmuje on jego uprzednią wiedzę o eksperymentowaniu — książkową i przypadkową, pogłoski o eksperymencie krążące w środowisku badanego, informacje przekazane mu w czasie werbowania do eksperymentu, wyposażenie laboratorium, wyjaśnienia i instrukcje udzielone przez eksperymentatora, osobę eksperymentatora, pretest, samą procedurę eksperymentalną — zachowanie się eksperymentatora, wyraźne i ukryte komunikowanie się w czasie trwania doświadczenia. Jedne z tych elementów, np. uprzednia wiedza, pomagają interpretować inne, np. instrukcje. Całość sygnałów pomagających badanemu domyślać się faktycznego celu doświadczenia nazwał Orne trudnym do przetłumaczenia terminem *demand characteristics of the experimental situation*[157], Badani aktywnie poszukują wszelkich wskazówek, które pomogłyby im zrozumieć, o co właściwie chodzi w sytuacji, w której się znajdują. Koniec końców, jeśli eksperymentator nie podejmuje specjalnych środków zapobiegawczych, wielu badanych łatwo domyśli się, do czego on zmierza.

W jaki sposób badany, który doszedł już do pewnej — trafnej lub nietrafnej interpretacji zamysłów eksperymentatora, uwzględni swoje odkrycie w wyborze

[157] Orne,op. cit., s. 237 - 8.

reakcji na bodźce, jakim został poddany? Zwróćmy najpierw uwagę, że pozytywne założenie tego pytania nie dla wszystkich badanych jest usprawiedliwione — na zachowanie pewnej części badanych ich własna interpretacja sytuacji może nie wywrzeć żadnego wpływu. Wystarczy jednak, że wpływ ten wystąpi u niektórych badanych, a już powstaje furtka dla błędu systematycznego.

Może się zdarzyć — zwłaszcza, gdy badanie jest przymusowe — że badany będzie zachowywał się „na złość" eksperymentatorowi wbrew imputowanym mu oczekiwaniom, będzie reagował w sposób przypadkowy, będzie się wygłupiał itd. Nie te rodzaje zachowania są jednak dominujące w eksperymentach psychospołecznych. Dominuje nastawienie pozytywne. Orne wykonał sugestywne doświadczenie, które pokazuje, do jakich poświęceń zdolni są badani w eksperymentach psychospołecznych[158]. Polecił on grupie badanych wykonywać pewne nudne i bezsensowne zadanie — dodawać liczby losowe na dostarczonych arkuszach; wypełnienie jednego arkusza wymagało 224 działań a każdy z badanych otrzymał 2000 takich arkuszy! Poinstruowawszy badanych i zabrawszy im zegarki eksperymentator wyszedł, zapowiadając swój „ewentualny" powrót. Po upływie 5 i pół godzin *eksperymentator* poddał się, a badani skłonni byli kontynuować swe dzieło dalej (jak się okazało, byli oni przekonani, że jest to sprawdzian ich wytrzymałości czy czegoś podobnego).

Jakież to motywy kierują zachowaniem się badanych w eksperymentach psychospołecznych? Dwa z nich wybijają się na czoło: chęć zyskania pozytywnej oceny ze strony eksperymentatora oraz swoiście rozumiana chęć przyczynienia się do postępu nauki; już widać, że nie są to cele sprzeczne — pierwszy z nich można osiągnąć za pośrednictwem drugiego.

Możliwość współwyznaczania zachowania badanych przez ich dążenie do zyskania uznania w oczach eksperymentatora jest wnioskiem z charakterystyki eksperymentu danej przez Rieckena: badany „usiłuje pokazać się w możliwie najlepszym świetle […] przez ukrycie lub wyeksponowanie, wyolbrzymienie lub pomniejszenie tych cech, które — jak sądzi — będą w eksperymencie ocenione dodatnio lub ujemnie"[159].

Zasugerowaną przez Rieckena interpretację rozwinął i udokumentował Milton Rosenberg[160] Wprowadził on będący już w szerokim obiegu termin „obawa o ocenę" *(evaluation apprehension)* na oznaczenie aktywnej, zabarwionej niepokojem troski badanego o zdobycie pozytywnej oceny eksperymentatora albo przynajmniej o niedostarczenie mu żadnych podstaw do oceny negatywnej". Osoby 'wrażliwe na swą ocenę, biorąc udział w badaniu, które ich zdaniem sprawdza nie tylko hipotezy, ale i ich samych, będą w większym stopniu niż osoby mniej wrażliwe na ocenę, aktywne w wykrywaniu na podstawie poszczególnych elementów procedury eksperymentalnej, jakie zachowanie zostanie uznane za normalne, dojrzałe, ujawniające kompetencję, inteligencję, przystosowanie itd. Osoby te będą starać się znaleźć sposoby zdobywania pozytywnej oceny eksperymentatora i unikania negatywnej. Ponieważ badani mogą oddzielnie dojść do sposobów podobnych, powstaje niebezpieczeństwo systematycznego zniekształcenia reakcji eksperymentalnych.

W celu wykazania wpływu tego mechanizmu na wyniki eksperymentu, Rosenberg wykonał pewne doświadczenie. Zakwestionował interpretację wyników serii eksperymentów nad zmianą postaw, które miały dowieść, że obrona poglądów sprzecznych z własnymi powoduje mniejszą ich zmianę wówczas, gdy nagroda za tę obronę jest wyższa. Zwolennicy teorii dysonansu poznawczego tłumaczyli ten wynik swoją teorią: łatwiej przyjąć poglądy sprzeczne, gdy są mało ważne a nagroda określa ich wagę. Rosenberg zinterpretował wyniki tych eksperymentów inaczej. Jego zdaniem, wysoka nagroda wzbudza podejrzenie badanego, że sprawdzana jest jego odporność na. zmianę postaw za pieniądze a także on sam.

[158] Ibidem.

[159] Riecken, op. cit., s. 31.

[160] Zob. M. J. Rosenberg, When dissonance fails: on eliminating evaluation apprehension from attitude measurement, w: Miller (ed.), op. cit., definicja ze s. 248.

Podejrzenie to wzmaga opór badanego a przynajmniej zmniejsza chęć dostarczenia danych świadczących o zmianie. Rosenberg zaprojektował więc eksperyment tak, że faza, w której następować mogło wzbudzenie podejrzeń i faza pomiaru postaw zostały oddzielone. Uzyskał wyniki różne od dotychczasowych: większa zmiana wystąpiła przy nagrodach wyższych — nagroda wzmacniała zmianę.

Zjawisko opisane przez Rieckena i Rosenberga jest dobrze znane w badaniach ankietowych: polega ono na udzielaniu przez respondenta takich odpowiedzi, które zyskałyby — jego zdaniem — aprobatę ankietera.

Interpretacja zachowań badanych w kategoriach dążenia do okazania badaczowi pomocy w rozwijaniu nauki związana jest z nazwiskiem Orne'a. W jego ujęciu — o czym już była mowa — motywem, który najczęściej skłania studentów do dobrowolnego udziału w eksperymentach, jest właśnie owo pragnienie przyczynienia się do postępu nauki. Jednakże z tym szlachetnym motywem wiąże się niestety zainteresowanie „sukcesem" eksperymentu, chęć okazania się „dobrym badanym", który „nie zepsuje" doświadczenia. „Zgoda, że badani zainteresowani są swym zachowaniem z punktu widzenia poprawy ich obrazu siebie samych — pisze Orne — niemniej jednak wydaje się, że nawet bardziej obchodzi ich użyteczność własnych zachowań. Moglibyśmy zatem słusznie oczekiwać, że badany w miarę swoich zdolności będzie w sytuacji eksperymentalnej zachowywał się w tak obmyślony sposób, by odegrać rolę «dobrego badanego» lub — innymi słowy — potwierdzić hipotezę eksperymentalną"[161], Rzecz jasna, jest to tylko pseudopotwierdzenie.

Również dla tej interpretacji dadzą się pomyśleć testy eksperymentalne. Im ważniejszy dla nauki wyda się badanym jakiś eksperyment, tym wyższa będzie ich motywacja do udziału w nim i tym częściej będą oni zachowywać się zgodnie z przypisywaną eksperymentatorowi hipotezą. Implikacji tych — o ile wiemy — nikt dotąd nie sprawdził.

Niektórzy badacze zajęli się natomiast porównaniem wyjaśniania zachowania badanych w terminach obawy o ocenę i chęci współpracy. Poszukiwali odpowiedzi na pytanie, czy badani zmieniają swe „naturalne" zachowanie dlatego, że chcą zadowolić eksperymentatora czy też dlatego, że próbują pokazać się w dobrym świetle.

Sigall, Aronson i Van Hoose[162] zauważyli, że między wyjaśnieniami Rosenberga i Orne'a nie ma sprzeczności tylko wtedy, gdy badani sądzą, że oceniani są ze względu na swą chęć współpracy. Jeśli natomiast oceniana ma być ich osobowość lub reakcje, to pojawia się możliwość konfliktu. W pewnych warunkach potwierdzenie hipotezy eksperymentatora wymagałoby ujawnienia ujemnie ocenianych aspektów osobowości, a w innych dla pokazania się od dobrej strony, konieczne byłoby zachowanie się wbrew hipotezie. Zdaniem autorów badani zainteresowani są raczej w swej ocenie niż we 'współpracy z eksperymentatorem; ten drugi motyw jest wtórny, jeśli w ogóle istnieje. Wysunęli oni przypuszczenie, że badani znając hipotezę eksperymentalną, nie będą usiłowali zachowywać się zgodnie z nią, jeśli miało by to stawiać ich w niekorzystnym świetle. Autorzy ci skonstruowali pewną prostą sytuację eksperymentalną, w której dwa dyskutowane motywy znalazły się w konflikcie. Wynik doświadczenia był zgodny z hipotezą. Dokonany przez nich przegląd literatury przytaczanej na poparcie tezy o sile dążenia do współpracy wskazuje także, że przytaczane tam wyniki badań mają małą wartość uzasadniającą.

Do podobnych wniosków doszedł Minor[163]. Badacz ten wykazał, że badani dostosowują się do uznanej *za* ważną hipotezy eksperymentalnej tylko wówczas, gdy sądzą, że ich zachowanie będzie podstawą do oceny ich osobowości. Badani potwierdzali hipotezę znacznie rzadziej, jeśli żywili przekonanie, że ich reakcje, choć

[161] Orne, op. cit., s. 237.

[162] H. Sigall, E. Aronson, Th. Van Hoose, The cooperative subject: myth or reality?, w: Miller (ed.), op. cit.

[163] M. W. Minor, Experimenter-expectancy effect as function of evaluation apprehension, „Journal of Personality and Social Psychology" 1970, vol. 15, nr 4.

istotne dla naukowych celów eksperymentu, nie będą podstawą oceny osobowości. Wyniki te skłaniają do rewizji tezy Orne'a o samodzielnym wpływie zainteresowania badanych postępem nauki na potwierdzenie hipotez. W wypadku prawdziwości tej tezy hipotezę potwierdzałyby również osoby nie dbające o swą ocenę, a one tego nie czynią.

Wyniki przytoczonych badań wskazują, że dążenie do uznania jest o wiele bardziej istotnym wyznacznikiem zachowania badanych niż chęć współpracy z eksperymentatorem w rozwijaniu nauki. Tezę tę popiera również pewna banalna refleksja. Mechanizm opisany przez Orne'a mógłby występować tylko w środowiskach, w których postęp nauki stanowi jedną z naczelnych wartości. Dążenie do uznania ze strony autorytetu jest — bądź co bądź — motywem mniej wyrafinowanym, a za to bardziej pospolitym.

Na sytuację eksperymentu można spojrzeć nie tylko od strony badanego — jak dotąd Czyniliśmy — lecz także od strony eksperymentatora. Jego osoba, postawy i zachowanie ważyć mogą istotnie na wynikach badania. Zjawisko wpływu eksperymentatora na wyniki eksperymentu znane jest pod nazwą „efektu eksperymentatora". Dostrzeżono je już dawno, ale w sposób systematyczny zbadał je dopiero Rosenthal w latach sześćdziesiątych[164].

Oddziaływania eksperymentatora na wyniki eksperymentu podzielił Rosenthal na takie, które nie zmieniają zachowania badanych i takie, które je zmieniają. Pierwsze polegają na stronniczej lub nieuczciwej obserwacji lub interpretacji zachowania i nie są osobliwe dla eksperymentu ani też szczególnie zajmujące. Ciekawe za to są wpływy drugiego rodzaju. Rosenthal wyróżnił kilka ich rodzajów.

Na zachowanie badanych wpływać mogą biopsychiczne cechy eksperymentatora, takie jak płeć czy wiek. Mogą one różnicować zachowania eksperymentatorów i tym samym badanych, mogą też różnicować postrzeganie eksperymentatorów przez badanych, którzy zareagują w sposób, jaki uznają za właściwy w świetle cech eksperymentatora. W obu wypadkach skutek jest ten sam.

Również cechy psychospołeczne i społeczne eksperymentatorów takie jak poziom lęku, potrzeba uznania, autorytaryzm, serdeczność, status społeczny mogą warunkować zachowanie badanych przez to, że wpływają uprzednio na zachowanie ich samych. Poziom lęku eksperymentatora zmierzony przed wejściem do laboratorium przejawia się w jego niespokojnym i ruchliwym zachowaniu w czasie eksperymentu i w nieco niepewnym tonie głosu. Badani o intensywnej potrzebie aprobaty dążą do uzyskania jej u eksperymentatora, z kolei eksperymentatorzy szukający uznania starają się o nie u badanych — powoduje to np. ich bardziej przyjazne zachowanie się wobec badanych.

Zachowanie eksperymentatora wobec badanych zależy też od pewnych czynników sytuacyjnych. Jakość wyposażenia laboratorium, jego mniej lub bardziej „naukowy" wygląd, doświadczenie badawcze eksperymentatora należą do tej klasy. Osobno wymieńmy rolę osoby trzeciej, jaką jest kierownik badań, jeśli eksperymenty prowadzi jego asystent. Może on nie zetknąć się w ogóle z badanymi a mimo to będzie miał wpływ na ich zachowanie. Samo istnienie kierownika, jego status i stosunki z eksperymentatorem będą kształtować zachowania tego ostatniego, a więc z kolei wpływać na zachowanie badanych.

Wymienione czynniki wchodzą ze sobą w najrozmaitsze interakcje. Istnieje też zjawisko sprzężenia zwrotnego między zachowaniem eksperymentatora i badanych: eksperymentator może reagować na zachowanie badanych, które przedtem sam wywołał.

Kolejnym rodzajem wpływów eksperymentatora na badanych są wpływy modelujące. Ich następstwem jest tendencja do osiągania przez badanych takich wyników w eksperymencie, jakie uzyskałby sam eksperymentator. Niestety empi-

[164] Referujemy niżej nast. prace tego autora: On the social psychology of the psychological experiment: the experimenter's hypothesis as unintended determinant of experimental results, „American Scientist" 1963, vol. 51, nr 2; Experimenter effects in behavioral research, New York 1976; Interpersonal expectations: effects of the experimenter's hypothesis, w: Rosenthal, Rosnow (eds.), op. cit.; Covert communication in psychological experiment, w: Miller (ed.), op. cit.

ryczne dowody tych tez pochodzą głównie z badań sondażowych (jedna z postaci efektu ankieterskiego) i klinicznych, nie ma jednak żadnego powodu, by przypuszczać, że nie może on wystąpić w badaniach eksperymentalnych.

Ostatnią kategorią oddziaływań eksperymentatora na zachowanie badanych są wpływy jego oczekiwań dotyczących wyników badań. Pierwsze doniesienie Rosenthala o tym zjawisku ma tytuł dobrze o nim informujący: „O społecznej psychologii eksperymentu psychologicznego: hipoteza eksperymentatora jako niezamierzony wyznacznik wyników eksperymentalnych". Jest to zjawisko frapujące — jeśli możemy sobie pozwolić na swobodne wartościowanie — i na nim zatrzymamy się najdłużej.

Otóż badacz przed rozpoczęciem badania formułuje pewne hipotezy co do jego wyniku a przynajmniej posiada jakieś oczekiwania w tym względzie. Oczekiwania te mogą wywierać istotny wpływ na zachowanie się badanych czyli na wynik doświadczenia. Działają one na rzecz potwierdzenia hipotezy. Hipoteza sformułowana przez eksperymentatora może więc działać jako samospełniającą się przepowiednia — samo jej postawienie podnosi szanse jej potwierdzenia. Przypomina się znane twierdzenie Thomasa: jeśli ludzie uznają jakieś sytuacje za rzeczywiste, to stają się one rzeczywiste w swoich konsekwencjach. Zjawisko to występuje częściej, gdy hipotezy sprawdzane zostały wymyślone przez samego eksperymentatora — są one drogie jego sercu. Siłę owego samo- potwierdzającego wpływu hipotezy różnicują ponadto takie czynniki jak płeć, status, doświadczenie badawcze eksperymentatora, czy jego skłonność do gestykulacji.

Nie należy tu sądzić, że eksperymentator musi komunikować badanym swoje hipotezy, a ci reagują w sposób, który ich zdaniem hipotezy te będzie potwierdzał. Nie byłoby to często możliwe choćby z tego powodu, że hipotezy te są skomplikowane i trudne do zrozumienia nawet po ich sformułowaniu. Badacz po prostu może komunikować oczekiwania dotyczące zachowania, które *jego* zdaniem hipotezy te potwierdzi.

Hipoteza może sama działać na swoje dobro bynajmniej nie tylko w eksperymentach laboratoryjnych. Zjawisko to bowiem zaobserwowano także w eksperymentach terenowych. Rosenthal i Jacobson przeprowadzili w szkole szeroko zakrojony eksperyment metodologiczny[165]. Poddali uczniów testowi inteligencji a następnie poinformowali nauczycieli, że był to test pozwalający przewidzieć rozwój intelektualny i wskazali im nazwiska wybranych losowo uczniów, którzy mieli rozwinąć się szczególnie intensywnie. Po upływie roku powtórny pomiar przekonał ich, że ci uczniowie, u których nauczyciele spodziewali się intelektualnego wzrostu, rzeczywiście wzrost ten osiągnęli. Tytuł książki, w której doświadczenie to opisano sam informuje o jego wyniku: *Pygmalion in the classroom.*

Co ciekawe, samokonfirujący wpływ hipotezy eksperymentatora na zachowanie badanych stwierdzono także w eksperymentach ze zwierzętami. Ba, może tam właśnie zauważono i opisano go najpierw.

Rosenthal i Fode polecili dwóm eksperymentatorom uczyć znajdowania wyjścia z labiryntu dwie grupy szczurów, wybrane losowo z jednorodnej kolonii[166], Przedtem jednakże poinformowali jednego z eksperymentatorów, że jego szczury są „bystre", a drugiego że jego szczury są „tępe". I cóż się stało? Szczury „bystre" rzeczywiście okazały się bystre i prędzej nauczyły się odnajdować drogę w labiryncie niż szczury „tępe".

Znacznie bardziej efektowny jest inny przypadek. Pod koniec ubiegłego wieku żył sobie w Niemczech koń imieniem Sprytny Jaś (Der *Fluge Hans).* Spryt owego konia polegał na tym, że potrafił wystukiwać kopytem prawidłowe odpowiedzi na stawiane mu pytania z zakresu arytmetyki i harmonii muzycznej. Spośród innych właścicieli „mądrych" zwierząt pan Sprytnego Jasia Herr von Olsen profesor gimnazjalny, wyróżniał się tym, że nie zabraniał badania tajemnicy cudownych

[165] R. Rosenthal, L. Jacobson, Pygmalion in the classroom: teacher expectation and pupils' intellectual development, New York 1968.

[166] R. Rosenthal, K. L. Fode, The effect of experimenter bias on the performance of the albino rat, „Behavioral Science" 1963, vol. 8, nr 2.

umiejętności swego podopiecznego. Skorzystał z tej okazji Otto Pfungst, psycholog. Zauważył on rychło, że gdy koń nie widział pytającego albo też sam pytający nie znał trafnej odpowiedzi na pytanie, tracił swój spryt. Widocznie pytający musiał dawać mu jakieś sygnały, kiedy zacząć i kiedy skończyć „odpowiedź". Bliższe badanie wykazało, że sygnałem do rozpoczęcia tupania było wychylenie głowy w stronę konia, nawet gdy nie było poprzedzone pytaniem. Sygnałami do zwolnienia i zaprzestania tupania było napięcie i rozluźnienie pojawiające się u pytającego w postaci ruchów brwi, nozdrzy, ruchów głowy, ważne były też nieuświadamiane sygnały głosowe. Koń nauczył się korzystać z tych wskazówek. To odkrywszy, Otto Pfungst sam stanął w roli konia. Kazał zadawać sobie pytania i wystukiwał odpowiedzi na nie palcem obserwując przy tym pytających. Z ich niekontrolowanych znaków w 23 przypadkach domyślił się trafnych odpowiedzi, a w 2 pozostałych pomylił się nieznacznie. Pfungst zbadał też, jakie cechy pytających podnoszą skuteczność warunkowania Sprytnego Jasia. Wskazał on tutaj na umiejętność obchodzenia się ze zwierzętami, spokojną władczość, koncentrację na prawidłowej odpowiedzi i oczekiwanie jej, skłonność do gestykulacji i wygląd świadczący o dobrym zdrowiu. Ciekawe jest to, że po dokonaniu koniecznych zmian, są to te same cechy, które sprzyjają sterowaniu badanymi w eksperymentach z ludźmi[167].

Casus Sprytnego Jasia wprowadza dobrze w następne zagadnienie. Jakimi to mianowicie kanałami dokonuje się pozastymulacyjny wpływ eksperymentatora na badanych, a zwłaszcza, w jaki sposób eksperymentator komunikuje badanym swe oczekiwania? I to pytanie doczekało się częściowych udokumentowanych wynikami badań odpowiedzi.

Prosty acz pomysłowy eksperyment Greenspoona[168] pokazuje, jak łatwo o wskazówki wzmacniające reakcje pożądane a wygaszające reakcje niepożądane, i to niezależnie od tego, czy eksperymentatorzy są ich świadomi i czy badani zdają sobie sprawę ze związku między ich zachowaniem a zachowaniem eksperymentatora. Greenspoon prosił badanych o wymienienie dowolnych rzeczowników i nagradzał zachęcającym pomrukiem wymienienie rzeczownika określonego rodzaju: uzyskał istotną różnicę w ilości nagrodzonych wyrazów między grupą kontrolną a eksperymentalną. Badani uczyli się sposobu reagowania z zachowania eksperymentatora.

Ale i sam eksperymentator uczy się w czasie eksperymentu, jak efektywniej uczyć badanych oczekiwanych zachowań. Jeśli badani po jakimś sygnale eksperymentatora nagradzają go zachowaniem zgodnym z jego hipotezą, to sygnał ten utrwala się. Zatem w eksperymentach, w których jeden badacz kontaktuje się kolejno z wieloma badanymi, badani w dalszej kolejności powinni częściej zachowywać się zgodnie z oczekiwaniami eksperymentatora niż badani w pierwszej kolejności. Rosenthal stwierdził, że tak rzeczywiście bywa.

Sygnałami komunikującymi badanym oczekiwania eksperymentatora i wzmacniającymi zachowania z tymi oczekiwaniami zgodne mogą być sygnały głosowe i behawioralne: „niechcący" wypowiadane słowa, dźwięki nieartykułowane, ton głosu, ruchy, mimika, uśmiechy itd. Ważne są szczególnie rozmaite formy komunikacji podprogowej, zwłaszcza jeśli ich nadawca jest w pewien sposób atrakcyjny. Są one szczególnie efektywne, zróżnicowane i trudne do wykrycia i kontroli. Dlatego to eksperymentatorzy dominujący, atrakcyjni dla badanych, łatwo nawiązujący z nimi osobisty kontakt, zachowujący się swobodnie lepiej sterują badanymi niż pozostali. Sprawy te nie są jednak jeszcze dokładnie zbadane. „Szkoda, że nie ma tu teraz Pfungsta, aby nam pomógł" — wzdycha Rosenthal[169].

Podobnie jak w analizie zachowania badanych, również w rozważaniach nad zachowaniem eksperymentatorów, trzeba postawić pytanie o zakres występowania opisanych procesów. Czy są one tylko wyjątkowo pojawiającymi się fenomenami czy też przeciwnie — są one plagą eksperymentalnych badań społecznych?

[167] Zob. O. Pfungst, Clever Hans (the horse of Mr. von Osten): a contribution to experimental, animal and human psychology (1907), New York 1965.
[168] Zob. Th. M. Newcomb i in., Psychologia społeczna, Warszawa 1970, s. 570.
[169] Rosenthal, Covert communication…, op. cit., s. 336.

Łatwo zgadnąć, że odpowiedź na takie pytanie jest niezmiernie trudna. Niektórzy krytycy prac Rosenthala zarzucali mu, że wyniki wielu swoich i cudzych eksperymentów interpretował życzliwie dla swych hipotez metodologicznych i sami bezskutecznie usiłowali wykazać wpływ hipotezy eksperymentatora na zachowanie badanych[170] Ważniejsze jest jednak to, że Rosenthal, a także inni badacze wykazywali w rozmaitych eksperymentach ów efekt z całą pewnością, że można go zademonstrować w pewnych eksperymentach na zawołanie, określić warunki jego występowania i wyjaśnić go teoretycznie. Efekt ten nie musi — rzecz jasna — występować w każdym badaniu, więc' niemożliwość jego wykazania w pewnych eksperymentach niczego do sprawy nie wnosi. Bliskie jest nam rozsądne stanowisko Levy'ego w tym względzie. Komentując krytykę prac Rosenthala przez Barbera i innych, autor ten pisze: „Wykazanie efektu eksperymentatora w doświadczeniu Rosenthala i Fode'a oraz w innych badaniach można porównać do odkrycia przez żonę dziwnego śladu szminki na kołnierzu mężowej koszuli. Bez względu na to, czy podejrzewała go ona o niewierność w przeszłości, czy też nie, wydaje się nieprawdopodobne, że jakakolwiek liczba nieskazitelnych kołnierzyków w przyszłości uwolni ją całkowicie od niepokoju o jego wierność małżeńską. Zatem wydaje się nieprawdopodobne, że pięć nieudanych prób odrzucenia hipotezy zerowej [o nieistnieniu efektu eksperymentatora] wykonanych przez Barbera i in. przekona wielu o rzadkości efektu eksperymentatora albo o tym, że nie musi on być brany pod uwagę w projektowaniu i interpretacji eksperymentów"[171]

Obydwie opisane wyżej tendencje tj. dążenie badanego do zyskania uznania w oczach eksperymentatora lub chęci okazania mu pomocy w rozwijaniu nauki oraz dążenie eksperymentatora do potwierdzenia się jego hipotezy uzupełniają się, by działać wspólnie na rzecz pseudopotwierdzenia sprawdzanej hipotezy. Tylko łączne wystąpienie tych tendencji wyjaśnia zjawisko samopotwierdzania: nadawane przez eksperymentatora informacje o tym, co jest dlań pożądane, natrafiają na skłonność badanego do czynienia tego właśnie. Cytowane już doświadczenie Minora pokazało, że gdy informacja o sprawdzanej hipotezie dotarła do osób sądzących, że ich zachowanie nie będzie podstawą oceny osobowości, wpływ hipotezy na zachowanie znikł.

Procesy psychospołeczne dokonujące się w sytuacji społecznej eksperymentu posiadają poważne znaczenie metodologiczne i pociągają za sobą takież następstwa dla badań eksperymentalnych. Tak więc:

1. To, czy sprawdzana w eksperymencie hipoteza zostanie przyjęta jako potwierdzona czy też zostanie odrzucona zależy nie tylko od jej przedmiotowej trafności. Zależy także od kontekstu psychospołecznego, w jakim zostaje sprawdzana: od całości sygnałów pozwalających badanym domyślać się jej a także i od ich postaw wobec eksperymentowania, a wreszcie od eksperymentatora samego, któremu zwykle jest miłe potwierdzenie się własnych przypuszczeń.

Jeśli badani nie uwierzą w podany im cel doświadczenia a odgadłszy właściwie cel faktyczny i związane z tym oczekiwania badacza zechcą zyskać jego uznanie lub mu „pomóc", to tym samym szanse potwierdzenia hipotezy rosną: orzekany przez hipotezę związek między zmiennymi może się wydać silniejszy niż jest faktycznie, a nawet może się zdarzyć, że stwierdzi się zależność tam, gdzie jej wcale nie ma.

Jeśli natomiast badani nietrafnie odgadną hipotezę eksperymentalną i do tej nietrafnej interpretacji dostosują swe zachowania, to ona a nie hipoteza sprawdzana może zostać pozornie potwierdzona. Gdy jednak hipotezy te są ze sobą sprzeczne, to szanse potwierdzenia się hipotezy eksperymentalnej maleją bez względu na to, ile hipoteza ta jest warta, nawet gdy jest prawdziwa. Eksperyment dowodzi wtedy związku słabego, podczas gdy ten jest silny albo wręcz ustala brak związku, podczas gdy ten faktycznie istnieje.

[170] Zob. Th. X. Barber, M. J. Silver, Fact, fiction, and the experimenter-bias effect, w: Miller (ed.), op. cit.

[171] Ibidem, wstęp A. G. Millera do części V, s. 320.

Kiedy jednak wziąć pod uwagę, że posiadana przez badanych praktyka lub wiedza dotycząca eksperymentu oraz kierujące sygnały eksperymentatora rozszerzają zakres występowania pierwszej z wyróżnionych sytuacji, można oczekiwać, że w psychologii społecznej udział pseudopotwierdzonych hipotez lub hipotez potwierdzonych słabiej niż zwykło się sądzić może być całkiem znaczny.

Wielu badaczy podjęło więc inspirowane przez te rozważania próby reinterpretacji lub powtarzania pewnych uznanych eksperymentów; Mills rozważył eksperyment Schachtera nad stosunkiem grupy do dewiantów i doszedł do wniosku, że niektóre jego wyniki z powodzeniem można wyjaśnić elementami interakcji między eksperymentatorem a badanymi[172]. Orne i Holland próbowali wyjaśnić zachowanie badanych w eksperymencie Milgrama nad posłuszeństwem przez odwołanie się do sygnałów sytuacyjnych wskazujących badanym, jak zachowywać się mają[173], Rosenberg — jak już informowaliśmy — powtórzył pewne eksperymenty nad zmianą postaw pod wpływem" perswazji i uzyskał z nich zupełnie różne od przyjętych wyniki, gdy wyeliminował wpływ podejrzliwości badanego na jego zachowanie. Page powtórzył eksperyment Staatsów nad uczeniem się znaczenia przez klasyczne warunkowanie i doszedł do przekonania, że wyniki tego doświadczenia lepiej wyjaśnić psychospołeczną teorią eksperymentu niż teorią Pawłowa[174]. Wreszcie, Schulman powtórzył eksperyment Ascha nad konformizmem i manipulując obecnością eksperymentatora pokazał, że reakcje badanych wyznaczone były nie tylko konformizmem wobec grupy, lecz także konformizmem wobec eksperymentatora[175]. Zbyteczne dodawać, że krytykowani nie zawsze akceptowali wyniki osiągnięte przez ich krytyków — wyniki te niejednokrotnie same są mocno niejednoznaczne.

Można powiedzieć, że samopotwierdzenie się hipotez równa się ustaleniu zależności pozornych, a samoobalenie się — ustalaniu pozornych niezależności. Zmienną kontaminującą jest tu czynnik psychospołeczny powiązany systematycznie z samą procedurą eksperymentalną. Z tego punktu widzenia skuteczna kontrola eksperymentalna okazuje się zadaniem trudniejszym niż mogłoby się to wydawać. Podjęta w celu usunięcia czynników zakłócających przebieg badanych zależności, wprowadza jednak na ich miejsce inne zmienne, których działanie powoduje ten sam skutek tj. trudność określenia roli zmiennej niezależnej w wywoływaniu zaobserwowanych zmian.

2. Komplikuje się powtarzalność eksperymentu i kontrola rzetelności jego wyników uzyskana tą drogą. Jeśli na wyniki eksperymentu mają wpływ szczególne cechy badanych, badacza i stosunków między nimi, to powtórzenie wymaga przy zestandaryzowaniu bodźców i pomiarów, różnicowania tych czynników. Dopiero, gdy różni badacze eksperymentując na różnych badanych uzyskają te same wyniki, będzie można powiedzieć, że wyniki te nie są obciążone wpływami związanymi z psychospołeczną sytuacją eksperymentu.

3. Rodzaj sytuacji, do których chciałoby się stosować wyniki eksperymentu nie musi się pokrywać z rodzajem sytuacji, w której wyniki te uzyskano. Wartość predyktywna hipotez sprawdzonych w warunkach sprzyjających ich potwierdzeniu dla warunków, w których czynniki sprzyjające nie działają, jest ograniczona — predykcje okazują się zawodne lub niedokładne. Jest to okoliczność druzgocąca, jeśli pamiętać o tym, że celem nauki jest opisanie i wyjaśnienie świata, a sam eksperyment jest tylko narzędziem do tego służącym.

Potrzeba uporania się z tymi problemami dała impuls badaniom nad eksperymentami, głównie także eksperymentalnym. Przy ich pomocy określa się wpływ rozmaitych okoliczności sytuacyjnych, proceduralnych i psychospołecznych na

[172] Th. M. Mills, A sleeper variable in small groups research: the experimenter, „Pacific Sociological Review" 1962, vol. 5, nr 1.

[173] M. T. Orne, Ch. H. Holland, On the ecological validity of laboratory deceptions, w: Miller (ed.), op. cit.

[174] M. M. Page, Social psychology of a classical conditioning of attitudes experiment, ibidem.

[175] G. J. Schulman, Asch conformity studies: conformity to experimenter and/or to the group?, „Sociometry" 1967, vol. 30, nr 1.

wyniki eksperymentu, konstruuje się nowe techniki badania i kontroli, wskazuje się optymalne warunki psychospołeczne ich stosowania. To, co przedstawiliśmy w tym punkcie było właśnie próbą poinformowania o niektórych wynikach tych badań. Pojawienie się tego typu badań jak tu omawiamy jest w historii nauk bardziej rozwiniętych regułą — eksperymentalne fragmenty nauk społecznych nie są od niej wyjątkiem.

Mógłby ktoś wyrazić wątpliwość co do wagi opisanych zjawisk, które, bądź co bądź, zaobserwowano tylko w jednym kraju — w Ameryce i w jednej kategorii społecznej — u studentów. To prawda, jednakże na eksperymentach wykonanych na amerykańskich studentach opiera się w bardzo ‘ dużym stopniu cała współczesna socjologia małych grup i psychologia społeczna, a przynajmniej ten dorobek tych nauk, z którym my tutaj mamy do czynienia. A jest oczywiście rzeczą niesłychanie interesującą, jakim modyfikacjom środowiskowym, a zwłaszcza kulturowym, poddana jest społeczna sytuacja eksperymentu. Niestety badań na ten temat nie znamy.

10.3. PROSTOTA SYTUACJI EKSPERYMENTALNYCH

Trzecim — obok osobliwości grup tworzonych na potrzeby eksperymentu i specyficznych interakcji społecznych zachodzących w czasie jego trwania — źródłem różnic między zachowaniami w sytuacjach eksperymentalnych i pozaeksperymentalnych jest uproszczony charakter pierwszych w porównaniu ze złożonym charakterem drugich.

Eksperyment kontrolowany polega na planowym poddawaniu badanych oddziaływaniom bodźców: zaczyna się od wprowadzenia czynników pojedynczych a potem stopniowo komplikuje się ich strukturę, czyni się to jednak tak, aby nie utracić możliwości ich kontroli. Jeśli zmiana eksperymentalna polega nie na wprowadzaniu czynników nowych, ale na wyłączeniu działania istniejących w danej sytuacji, to dodatkowo upraszcza się sama badana sytuacja, również w sposób stopniowy i kontrolowany. W sytuacjach życiowych natomiast ludzie reagują zwykle na złożone i wzajemnie powiązane konfiguracje czynników a jeśli zdarza się, że działanie jednego czynnika jest szczególnie wyraźne dzięki stabilizacji lub wyłączeniu się innych, to badacze wykazują zaraz skłonność do nazywania tych sytuacji quasi-eksperymentalnymi właśnie.

10.4. CZY SYTUACJE EKSPERYMENTALNE SĄ SZTUCZNE?

Uświadomienie sobie przedyskutowanych różnic między sytuacjami eksperymentalnymi a nieeksperymentalnymi prowadzi do potocznego poglądu, że sytuacje pierwszego rodzaju są „sztuczne” ’i że wobec tego ich badanie niewiele może pomóc w wyjaśnianiu i przewidywaniu zjawisk „rzeczywistych”. Powiada się, że badania eksperymentalne nie tyle odsłaniają rzeczywistość, co ją tworzą, a zamiast odkrywać fakty, ujawniają tylko artefakty — zjawiska wywołane przez badanie, mylnie brane za istniejące niezależnie od badania. Często też formułuje się dylemat: „eksperyment czy rzeczywistość?” Dokładniejsze przyjrzenie się tym wypowiedziom może jednak ujawnić, że mimo niekwestionowanych różnic między sytuacjami eksperymentalnymi, ani zarzuty nie są tak oczywiście trafne, ani odpowiedź na dylemat tak jednoznaczna, jak to sugeruje kategoryczność i elegancja, z jaką je sformułowano.

Przystępując do rozpatrzenia pierwszego powodu odmienności zachowań w sytuacjach eksperymentalnych i nieeksperymentalnych tj. specyfiki pewnego typu grup biorących udział w eksperymencie, przypomnieć wypada, że zarzut sztucz-

ności nie musi się stosować do eksperymentów prowadzonych na grupach naturalnych, istniejących niezależnie od badania. Jeśli natomiast odniesiemy go do grup tworzonych doraźnie, to nasuwa się uwaga, że znane są liczne eksperymenty na takich grupach przeprowadzone, które dostarczyły wiedzy o rzeczywistości pozalaboratoryjnej, której nie odrzuca się jako artefaktualnej, a przeciwnie, darzy się respektem za pogłębienie rozumienia ludzi i świata. Nikt rozsądny nie utrzymuje, że pewne działy socjologii małych grup czy psychologii społecznej są bezwartościowe tylko dlatego, że oparte są na eksperymentach laboratoryjnych. Skoro tak, to widocznie sam fakt doraźnego charakteru grup w eksperymencie niczego jeszcze nie przesądza. Na pewno sprawa jest bardziej skomplikowana.

Nie ulega wątpliwości, że między grupami naturalnymi a doraźnymi grupami laboratoryjnymi, obok poważnych a wskazanych wyżej różnic, istnieją także równie ważne podobieństwa. Pochodzą one z dwóch przynajmniej źródeł. Po pierwsze, grupy obu rodzajów zbudowane są z tych samych elementów tj. ludzi, wyposażonych w dyspozycje, które nie znikają wraz z przekroczeniem przez nich progu laboratorium. Po wtóre, związki między elementami tych grup mają w obu wypadkach charakter społeczny. Zatem zachowania eksperymentalne są sztuczne w takim stopniu, w jakim podporządkowane są tym aspektom funkcjonowania grup eksperymentalnych, które nie mają swych odpowiedników w życiu grup naturalnych. Są zaś naturalne o tyle, o ile wyznaczone są aspektami życia społecznego wspólnymi dla obu grup.

Istotnie więc w grupach powołanych do życia *ad hoc,* nie można badać roli utrwalonych w grupie stosunków społecznych w kierowaniu zachowaniem, ani też nie można badać takich aspektów życia grupy, które od owych stosunków zależą. Wyniki takich badań mielibyśmy prawo uznać za sztuczne. Można jednak z powodzeniem eksperymentować nad — powiedzmy — społecznymi uwarunkowaniami postrzegania, gdyż w tej dziedzinie oddziaływanie sytuacji społecznych na zjawiska badane nie ma charakteru aktualnego, lecz jest utrwalone w dyspozycjach jednostek. Eksperyment Brunera i Goodmana nad postrzeganiem monet przez dzieci wywodzące się z różnych klas społecznych jest realizacją tej możliwości[176]. Dalej, można badać takie zachowania, kształtowane przez sytuacje aktualne, których trwałość jest bez znaczenia. Badania nacisku grupowego, perswazji, konformizmu, władzy mieszczą się w tej kategorii. Wreszcie, w wielu eksperymentach istnienie zastanych struktur byłoby tylko przeszkodą. W badaniach skuteczności działania grup zadaniowych zorganizowanych w różny sposób trzeba właśnie uwolnić się od struktur zastanych i organizować grupy zadaniowe od podstaw. Na pewno nie wymieniliśmy wszystkich sytuacji, w których zarzut sztuczności uzasadniany doraźnym charakterem grup eksperymentalnych nie byłby na miejscu.

Zachowanie badanych w laboratorium zależy również od stopnia realizmu eksperymentu. Eksperyment jest realistyczny, „jeśli sytuacja eksperymentalna jest realistyczna dla badanego, jeśli go wciąga, jeśli jest on zmuszony brać ją poważnie, jeśli ma ona na niego wpływ". Ten rodzaj realizmu Aronson i Carlsmith nazywają realizmem eksperymentalnym w odróżnieniu od realizmu życiowego *(mundane),* który mierzy się „prawdopodobieństwem wystąpienia w «świecie rzeczywistym» zdarzeń pojawiających się w laboratorium"[177]

Wiemy już, że wiele momentów charakteryzujących laboratorium, choćby samo miejsce i niecodzienny charakter zadań, niebywale utrudnia osiągnięcie realizmu eksperymentalnego. Stąd stopień jego bywa często niski. Nie ma jednakże zasadniczych przeszkód, które uniemożliwiłyby jego zwiększenie. Osiągnięcie wysokiego stopnia realizmu eksperymentalnego zależy w ogromnej mierze od pomysłowości eksperymentatora, od jego objaśnień, scenariusza, procedury itd. Stąd obok badań krytykowanych — zresztą zwykle *ex post* — za brak realizmu ekspe-

[176] Zob. S. Mika, Wstęp do psychologii społecznej, Warszawa 1972, s. 42-4.

[177] E. Aronson, J. M. Carlsmith, Experimentation in social psychology, w: G. Lindzey, E. Aronson, (eds.), The handbook of social psychology, Reading, vol. 2, Mass. 1988, s. 22.

rymentalnego, istnieją doświadczenia, w których udało się osiągnąć jego bardzo wysoki poziom. I znowu, badanie Milgrama nad posłuszeństwem będzie dobrym przykładem. W czasie wywiadu poeksperymentalnego 56% badanych stwierdziło, że było całkowicie pewnych, że stosuje wstrząsy elektryczne a 24% miało pewne wątpliwości, lecz sądziło, że prawdopodobnie wstrząsy stosuje. Również objawy skrajnego napięcia psychicznego u wielu badanych świadczyły, że sytuację biorą oni na serio[178].

Jeśli uda się sprawić, że sytuacja eksperymentalna będzie realistyczna dla badanych, to wcale nie jest konieczne, by była ona w dodatku podobna do sytuacji codziennych. (Rzecz jasna, czasami realizm drugiego rodzaju jest jedynym sposobem zapewnienia realizmu pierwszego rodzaju). Trzeba bowiem mieć na uwadze, że sytuacje eksperymentalne, a zwłaszcza wprowadzone bodźce obmyślane są tak, by stanowiły trafne operacjonalizacje zmiennych niezależnych występujących w sprawdzanych hipotezach teoretycznych, a nie tak, by naśladowały życie codzienne. Do przewidywania zachowań w warunkach naturalnych posłużą potwierdzone już hipotezy teoretyczne a o ich odpowiednikach operacyjnych i tak się zapomni. Przed zastosowaniem ich trzeba będzie zresztą wskazać fakty pod nie podpadające — mogą one w małym stopniu przypominać zdarzenia w laboratorium. Gdyby miało być inaczej, tzn. gdyby celem eksperymentu było naśladowanie sytuacji naturalnych, to po cóż z takim trudem tworzyłoby się w laboratorium to, co poza nim jest łatwo dostępne — prościej byłoby wszak obserwować gotowe sytuacje naturalne! Widocznie w życiu nie sposób znaleźć takich operacjonalizacji zmiennych teoretycznych, jakie można stworzyć w laboratorium.

Podsumujmy tę partię rozważań. Sztuczność względnie naturalność sytuacji i zachowań w eksperymentach laboratoryjnych nie zależy tylko od miejsca badania, lecz w dużym stopniu także od umiejętności selekcji dziedziny badaniami od pomysłowości eksperymentatora. Istotne jest też, że podobieństwo sytuacji eksperymentalnych do sytuacji naturalnych samo w sobie jest bez wartości a przeciwnie, dopiero sytuacje niecodzienne są ważne dla sprawdzania hipotez.

Druga grupa argumentów na rzecz tezy o sztuczności zachowań eksperymentalnych odwołuje się do opisanych wyżej osobliwości społecznej sytuacji eksperymentu, a zwłaszcza do stosunków społecznych zawiązujących się między badaczem a badanymi, które mogą prowadzić do pseudokonfirmacji sprawdzanych hipotez. Zarzut sztuczności poparty przykładami wpływów oczekiwań eksperymentatora na zachowanie badanych, i dostosowywania się badanych do tych oczekiwań po to, by zyskać uznanie eksperymentatora lub „pomóc” mu, wydaje się w pierwszej chwili słuszny w stopniu nie pozostawiającym miejsca na żadne wątpliwości. Wszak badania, w trakcie których dzieją się takie rzeczy, służą właśnie raczej produkowaniu artefaktów niż odsłanianiu rzeczywistości. Nikt też nie zaprzeczy, że w sytuacjach życiowych te kłopotliwe zjawiska nie występują, wobec czego przenoszenie wyników eksperymentalnych na warunki rzeczywiste jest ryzykowne. Wszystko to prawda, jednakże poprzestanie na takich sformułowaniach byłoby niedopuszczalnym uproszczeniem.

Ujawnienie, że wyniki badania podejrzane o artefaktualność faktycznie mają taki charakter jest równoznaczne z radykalną zmianą ich statusu metodologicznego. To, co się brało za naturalne, okazało się tylko proceduropochodne. Został wykryty błąd, a więc uczyniono ważny krok w poznaniu rzeczywistości. Błąd w nauce jest rzeczą normalną, nie ma nauk, których pomiary nie są nim obciążone, są tylko nauki, które swych błędów nie dostrzegają; może nawet stopień rozwoju nauki wyraża się w łatwości wykrywania przez nią własnych błędów. Złem dopiero jest błąd nierozpoznany jako błąd.

Teraz sprawa jest już prostsza. Artefaktualny charakter wyników eksperymentalnych można wykryć przez prowadzenie badań nad badaniami: ustalenie, jaka część zaobserwowanych wyników ma charakter metodopochodny nie pozwoli

[178] Zob. S. Milgram, Interpreting obedience; error and evidence, w: The individual in a social world, New York 1978.

stwierdzić, że wyniki te są po prostu sztuczne. Można też obmyślać i stosować takie techniki eksperymentalne, dzięki którym wpływ samego badania na badanych udaje się minimalizować i pomierzyć. Wyniki takich badań nad badaniami już referowaliśmy i zapewne stworzyliśmy niepożądane wrażenie, że wiedza uzyskana w drodze eksperymentalnej jest niewiele warta. Kolej więc na przedstawienie metod, które pozwalają całkiem dobrze sobie z dyskutowanymi wpływami radzić. Przekonamy się, że eksperymentator nie jest wobec nich bezradny[179].

Fakt kierowania się badanych domniemywaną przez siebie hipotezą eksperymentalną zmusza do ukrywania przed nimi celu eksperymentu.

Można nie podawać im żadnej hipotezy, lecz nie jest to rozwiązanie dobre, gdyż badani wymyślają wówczas hipotezy własne i przypisują je eksperymentatorowi; ich poszukiwania mogą pójść w tym samym kierunku i doprowadzić do odgadnięcia hipotezy sprawdzanej. Lepszym wyjściem jest więc podawanie im fałszywego celu doświadczenia — chce się przez to zmylić badanych a zniekształceniom nadać kierunek znany. Ale niektórzy badani dysponują już pewną wiedzą o eksperymentach i nie biorą informacji eksperymentatora dosłownie. „Mogą oni — pisze Kelman o studentach amerykańskich — nie znać dokładnego celu poszczególnego eksperymentu, w którym uczestniczą, ale zwykle wiedzą przynajmniej, że nie jest nim to, co mówi eksperymentator”[180]. W każdym razie jednak podanie fałszywej hipotezy rozprasza poszukiwania badanych i zmniejsza niebezpieczeństwo błędu systematycznego. Podejrzliwy badany będzie spekulował, lecz oddali się raczej od hipotezy prawdziwej niż się do niej przybliży.

Oddzielenie pomiarów i manipulacji zmniejsza wpływ postaw badanych wobec eksperymentu na ich zachowanie. Badani sądzą, że biorą udział w dwóch eksperymentach — będzie to dla nich łatwiejsze, jeśli każde ze spotkań odbędzie się w innym miejscu. Tak postępował Rosenberg w celu wykrycia wpływu obawy czy uzyska się właściwą ocenę na wielkość zmiany postaw. Następni badacze mogli już planować swe badania tak, by tego wpływu uniknąć.

Można też przekonać badanych, że nie występują w roli badanych, lecz w roli eksperymentatorów. Milgram tak właśnie zrobił.

Eksperyment można zaaranżować poza laboratorium i to tak, że badany nie domyśli się w ogóle, że jest badanym. Tak postąpili np. Abelson i Miller w badaniach wpływu negatywnej i obraźliwej perswazji na zmianę poglądów[181]. Eksperymentator występujący w roli ankietera rozpoczynał wywiad z osobą siedzącą na ławce w parku, po czym siedzący niby przypadkowo na tej samej ławce osobnik, faktycznie — współpracownik eksperymentatora, wygłaszał przeciwne poglądy, przy czym bezlitośnie wyśmiewał badanego. Wówczas eksperymentator pytał badanego po raz wtóry.

Użycie behawioralnych a nie werbalnych miar zmiennej zależnej także zmniejsza wpływ postawy badanego wobec eksperymentu na jego reakcje. Chęć pomocy eksperymentatorowi lub zrobienie na nim dobrego wrażenia łatwiej wyraża się w odpowiedziach niż w zachowaniach.

Odpowiednio dobrany schemat eksperymentu również jest istotny. Polecane są zwłaszcza schematy z jednokrotnym pomiarem, gdyż dopuszczenie pretestu równa się dodaniu jeszcze jednego źródła informacji o hipotezie eksperymentatora. „Schemat z dwukrotnym pomiarem jest schematem, który często komunikuje hipotezę badanemu. Tak np. jeśli poprosi się badanego o określenie swej postawy w pewnej kwestii, a następnie o przeczytanie przekazu dotyczącego tej kwestii, a w końcu o ponowne określenie swej postawy w tej samej sprawie, to oczywiste wyda mu się oczekiwanie, że manipulacja zmieni w pewien sposób jego postawę” — piszą Aronson i Carlsmith[182]. O zagadnieniach doboru odpowiednich schematów tak, by zmniejszyć uwrażliwiające wpływy pretestu mówiliśmy obszerniej we właściwym miejscu.

[179] Przegląd tych technik opiera się zasadniczo na studium Aronsona i Carlsmitha, cyt. w przyp. 33.
[180] H. C. Kelman, Time to speak: on human values and social research, San Francisco 1968.
[181] Cyt. za Aronsonem i Carlsmithem, s. 46.
[182] Ibidem, s. 66.

Obok technik minimalizujących wpływ *demand characteristics* na zachowanie badanych istnieją również techniki jego pomiaru.

W czasie wywiadu poeksperymentalnego można pytać każdego z badanych o jego hipotezę, a gdyby się okazało, że zachowania poszczególnych badanych korelują się mocniej z ich własnymi hipotezami niż z hipotezą testową, byłby to dowód siły tendencji do podporządkowania zachowania hipotezie. W czasie wywiadu przedeksperymentalnego — proponuje Orne — można wyjaśnić losowo wybranej grupie badanych cel eksperymentu, wykonać na niej protest, poddać ją manipulacji i zamiast zmierzyć ją ponownie, zapytać o co — ich zdaniem — chodzi w eksperymencie, jakich zachowań eksperymentator od nich oczekuje itp., by w końcu porównać opis zachowań uznanych za oczekiwane z pomiarem zachowań reszty badanych[183]. Również niektóre schematy eksperymentalne skonstruowano specjalnie po to, by mierzyć owe wpływy. Mowa głównie o schemacie Solomona.

Również efekt eksperymentatora można poważnie zmniejszać za pomocą pewnych zabiegów[184].

Można prowadzić eksperymenty tak, by eksperymentator mający bezpośredni kontakt z badanymi nie znał hipotez. Jest to jednak metoda nieefektywna, gdyż może on jeszcze łatwiej niż badani domyślić się celu przedsięwzięcia.

Inny, lepszy sposób polega na tym, że eksperymentator wie, jaka hipoteza jest sprawdzana, lecz posiada ograniczoną wiedzę o tym, którzy badani otrzymali poszczególne warianty bodźca. Wiedza ta może być ograniczona przez opóźnienie przydziału poszczególnych grup do kolejnych wariantów aż do momentu manipulacji. W eksperymentach czynnikowych każdy z kilku eksperymentatorów zna przydział osób tylko do jednego czynnika.

Ograniczony może być też kontakt eksperymentatora z badanymi. Można to sprawić przez zastosowanie jakichś środków mechanicznych, np. standardowych instrukcji nagranych na taśmę magnetofonową. Ciekawszą metodą jest choćby podział eksperymentu na wielu eksperymentatorów, z których każdy wykona kolejną fazę badania. Taki podział utrudni badanym nauczenie się sygnałów eksperymentatorów, a eksperymentatorom utrudni nauczenie się komunikowania badanym swoich oczekiwań.

W celu pomiaru siły wpływu oczekiwań eksperymentatora na zachowanie badanych można wykonać specjalny eksperyment, w którym z bodźcem i jego brakiem zostanie skrzyżowane oczekiwanie i jego brak.

Próbą eliminacji omawianych wpływów jest również większy udział w eksperymentach badanych pochodzących ze środowisk nie obeznanych z wiedzą psychospołeczną i praktyką eksperymentalną. Tacy badani z pewnością nie będą umieli łatwo odgadnąć intencji eksperymentatora. Pochodząc ze środowisk, w których postępu nauki nie ceni się tak wysoko jak wśród studentów-ochotników, a także mniejszy jest autorytet eksperymentatora, badani nie, będą skłonni do okazywania mu pomocy ani do zabiegów o jego uznanie.

Podsumujmy ten punkt: zarzut sztuczności sytuacji eksperymentalnych uzasadniany interakcjami społecznymi w czasie badania wpływającymi ujemnie na wartość wyników traci na zasadności w miarę poznawania i oznaczania tych wpływów oraz w miarę rozwoju technik eksperymentalnych, które pozwalają minimalizować, a przynajmniej kontrolować owe wpływy. Zniekształcenia wyników z nich wypływających w żadnym razie nie są czymś koniecznym dla eksperymentalistyki psychospołecznej.

Pozostaje trzeci rodzaj krytyki eksperymentów za ich sztuczność: wskazywanie na uproszczony charakter sytuacji stworzonych w porównaniu ze złożoną naturą sytuacji spotykanych w życiu codziennym. Mówienie jednak o sytuacjach eksperymentalnych jako sztucznych i uproszczonych w żadnym wypadku nie powinno być krytyką, jest bowiem jedynie wypowiadaniem pleonazmów. Pierwsze określenie ma charakter definicyjny: każda sytuacja eksperymentalna jest sztuczna

[183] Orne, op. cit.

[184] Zob. Aronson, Carlsmith, op. cit. Rosenthal, Experimenter effects in behavioral research, op. cit.

w tym sensie, że badacz ją stworzył a nie wynalazł. Druga własność jest również konieczną, choć pozadefinicyjną cechą eksperymentu a ujemne jej wartościowanie dowodzi jedynie braku zrozumienia czym eksperyment jest a czym nie jest. Eksperyment nie jest kopiowaniem zjawisk, jakie zdarzają się w życiu. Sytuacje naturalne są na ogół zbyt skomplikowane, by można było ocenić skutki działania poszczególnych czynników i osiągnąć w tym względzie rozstrzygające wyniki. Eksperyment jest natomiast świadomym i zamierzonym uproszczeniem sytuacji życiowych i to właśnie dzięki temu jego wyniki posiadają tak dużą wartość dowodową i doniosłość dla budowy teorii. Sprawy tej dotknęliśmy w rozdziale drugim — obecnie przyjdzie nam rozważyć ją w sposób nieco bardziej systematyczny.

W ostatnich latach wyraźnie zwrócono uwagę na fakt, że twierdzenia i teorie w wielu dziedzinach nauk przyrodniczych oraz w rozwiniętych gałęziach nauk społecznych mają charakter idealizujący. Sformułowano także postulat nawołujący do śmielszego niż dotąd stosowania idealizacyjnej strategii budowy teorii na terenie zjawisk społecznych.

W niektórych naukach empirycznych twierdzenia formułuje się jako ważne dla pewnych warunków idealnych, sytuacji granicznych, przypadków czystych. Rodzaje owych przedmiotów abstrakcyjnych określane są w poprzednikach tych twierdzeń a znajdowane w rzeczywistości sytuacje faktyczne jedynie się do nich mniej lub bardziej przybliżają. To zaś, co głoszą następniki jest — jako następstwo odległości między postulowanymi warunkami idealnymi a warunkami rzeczywistymi — różne od wyników obserwacji tych drugich. Teoria jakiejś dziedziny zjawisk jest systemem twierdzeń idealizujących — orzekają one o zależnościach między rozmaitymi cechami przedmiotów abstrakcyjnych. Twierdzenia w obrębie teorii dają się porządkować według stopnia abstrakcji tj. ilości pominiętych względnie przejaskrawionych cech przedmiotów rzeczywistych. Budowa teorii rozumiana jako proces polega na budowie zrazu bardzo abstrakcyjnego, uwzględniającego tylko kilka ściśle określonych zmiennych, modelu właściwej dziedziny zjawisk i sformułowaniu twierdzeń o związkach między tymi zmiennymi przy założeniu ich izolacji od działania zmiennych pominiętych. Następnie ów model uszczegóławia się, wzbogaca przez włączenie doń kolejnych, uprzednio pominiętych zmiennych oraz stopniowo modyfikuje się przedtem sformułowane twierdzenia, tak by uwzględniały działanie tych zmiennych. W ten sposób dochodzi się aż do poziomu, na którym rozbieżności między sytuacjami obserwowanymi a konstruowanymi będzie można zlekceważyć bez istotnej szkody dla wytkniętego celu poznawczego. Bez odwołania się do tego celu optymalny stopień abstrakcji modelu i wiedzy o nim nie może być rozsądnie określony, gdyż z jednej strony winny one być abstrakcyjne i proste, by móc być operatywnymi, a z drugiej realistyczne i skomplikowane, by móc być Wiernymi — wymiary te nie mają skali wspólnej.

W rozwiniętej postaci podejście idealizujące jest w zakresie nauk społecznych trudniej stosowalne niż gdzie indziej, aczkolwiek i tu istnieją godne uwagi przykłady jego zastosowania w postaci teorii ekonomicznych, psychologicznych czy nawet socjologicznych sformułowanych dla pewnych modeli. Jednakże niezależnie od tego, idealizacyjny sposób formułowania twierdzeń występuje w pospolitych formach zalążkowych. Mowa o sytuacjach, w których wygłasza się zdania, że w pewnych warunkach jakaś zależność występuje wyraźniej niż w innych, że raz jest czysta a drugi raz zakłócona, że pewien związek daje się obserwować, gdy nie zachodzą określone warunki przeszkadzające a wobec pojawienia się ich — znika, że jakaś zależność ma swoje stale modyfikatory a pewne zmienne kontaminują określony rodzaj zależności, że syndromatyczna natura zjawisk nie pozwala śledzić związków między poszczególnymi zmiennymi, że zależności między zjawiskami obserwowalnymi w rzeczywistości mają nieuchronnie charakter statystyczny, że badacz formułując jakieś prawo, abstrahuje od pewnych istotnych czynników; mowa wreszcie o sytuacjach, w których wygłaszane twierdzenie opatruje się klauzulą *ceteris paribus*.

We wszystkich tych wypadkach wypowiadający wskazane zdania odwołuje się milcząco lub wyraźnie do pojęcia twierdzenia idealizacyjnego, które obowiązuje

ściśle a nie w przybliżeniu jedynie i dopiero w warunkach nierzeczywistych; polegają one na pominięciu. wyłączeniu zmiennych zakłócających. Warunki te ponadto wartościuje się dodatnio, uznaje się za pożądane z punktu widzenia skuteczności poczynań naukowych. A celem eksperymentu jest nie co innego jak tylko tworzenie i mnożenie sytuacji możliwie najbliższych owym sytuacjom pożądanym. Eksperymentator wyłącza działanie pewnych zmiennych, przez co tworzy warunki umożliwiające obserwację zależności w postaci możliwie czystej, nie zakłócanej przez te czynniki. Eksperymentator wprowadza nowe czynniki tak, by móc ustalić zachowanie się stwierdzonych zależności w warunkach oddziaływania zróżnicowanych modyfikatorów i ich precyzyjnie kontrolowanych kombinacji. Ten drugi moment wart jest szczególnego podkreślenia, gdyż celem badaczy jest formułowanie zdań pozwalających na przewidywanie zjawisk rzeczywistych — dochodzenie do twierdzeń ważnych tylko dla warunków idealnych, czystych jest jedynie krokiem do tego celu wiodącym.

Od eksperymentów z zasady nie należy więc oczekiwać ścisłych opisów procesów życiowych, lecz jedynie dobrych pierwszych przybliżeń do takich odwzorowań. Warunki, które różnią sytuacje eksperymentalne od sytuacji spotykanych w życiu mogą okazać się przynajmniej modyfikatorami siły, a ponadto kształtu czy nawet kierunku zależności ustalonych eksperymentalnie. Za to warunki owe powinny być wyraźnie wyspecyfikowane, chyba że zgodnie uznaje się je za oczywiste. Kiedy nie są one podane, to wartość twierdzenia wiążącego bodźce z reakcjami jest bardzo mała. Twierdzenie to ma dużą wartość systematyczną tylko na pozór — w istocie jest trudne do obalenia nie z racji swej trafności, ale z powodu braku precyzji — można go bronić bez końca zawężając ustawicznie *ad hoc* zakres jego stosowalności tak, aby ten nie obejmował kontrprzykładu.

Zakaz mechanicznego przenoszenia wyników badań eksperymentalnych na sytuacje obserwowane w życiu nie jest równoznaczny z usprawiedliwieniem nazywania zachowań eksperymentalnych sztucznymi w sensie „nierzeczywistymi” i lekceważenia podobieństw między nimi a zachowaniami w warunkach naturalnych. Eksperyment ma swoją rzeczywistość i jest ona nie mniej naturalna niż rzeczywistość po prostu. Laboratorium jest miejscem rzeczywistym, a ludzie w nim tak czy inaczej się zachowujący są także realnymi postaciami. Stąd wiedza o zachowaniach laboratoryjnych jest wiedzą o realnych ludzkich osobowościach i realnych ludzkich zachowaniach. Stąd przeciwstawieniem sytuacji eksperymentalnych nie są — ściśle mówiąc — sytuacje rzeczywiste, ale codzienne.

*

* *

Dylemat „eksperyment czy rzeczywistość?” jest zatem alternatywą mylącą. Problemem jest nie „czy?”, ale „jak?”, „z czego na co?” ekstrapolować. Rzecz sprowadza się przeto do ustalenia empirycznych relacji między sytuacjami eksperymentalnymi a życiowymi oraz posiadania teorii usprawiedliwiającej ekstrapolacje. Niedorozwój wiedzy pozwalającej przechodzić z poziomu sytuacji eksperymentalnych na poziom sytuacji życiowych jest tą okolicznością, która w wielkim stopniu jest odpowiedzialna za poczucie sztuczności sytuacji eksperymentalnych w badaniach społecznych.

CZĘŚĆ IV REAKCJE NA PRZESZKODY. EKSPERYMENTY ZASTĘPCZE NA MODELACH

Dążenia badań społecznych do ustalania związków przyczynowych napotykają na utrudniające ich realizację przeszkody. Metoda eksperymentalna, najlepiej do tego celu przydatna i dlatego szeroko a owocnie stosowana przez przyrodoznawstwo, na ich własnym gruncie posiada łatwo uchwytne i dosyć restryktywne granice stosowania. Na przeszkodę stojącą na drodze do celu można zareagować trojako — rezygnacją z celu, usunięciem przeszkody lub znalezieniem innej drogi do celu.

W dziedzinie rozważanej pierwszy rodzaj reakcji na przeszkody reprezentowany jest przez rezygnację z uprawiania nauki bądź odmawianie miana nauki nieeksperymentalnym dziedzinom nauk społecznych Źródłem obu możliwych postaw jest utożsamienie metody eksperymentalnej z metodą naukową: jeśli okaże się, że eksperyment może być stosowany w zakresie dosyć wąskim, to reakcje te stają się zrozumiałe. Trudno o konkretne ich przykłady — nie mają one charakteru demonstracyjnego wycofania się z nauki z powodu nagłego ujawnienia się przeszkód dla eksperymentowania, lecz polegają raczej na nieangażowaniu się w ogóle w badania społeczne, gdyż granice eksperymentu są z góry dobrze znane, a większe możliwości stosowania metod ścisłych, „prawdziwie naukowych" oferują nauki przyrodnicze.

Nauka jest dziedziną aktywności w wysokim stopniu racjonalnej, więc przeważają w niej dwa pozostałe rodzaje reakcji na przeszkody. Jeśli badania społeczne nie mogą osiągnąć jednego z najważniejszych swych celów tj. ustalania twierdzeń przyczynowych za pomocą eksperymentu, narzędzia wybornego, lecz trudnego do zastosowania, to mając na widoku ów cel, starają się pokonać trudności eksperymentowania lub posłużyć się innymi metodami doń prowadzącymi.

Pokonywanie przeszkód utrudniających eksperymentowanie polega na konstruowaniu i stosowaniu nowych schematów i procedur eksperymentalnych oraz na ustalaniu najwłaściwszych warunków ich wykorzystania. Znaczne fragmenty niniejszej pracy poświęcone były właśnie temu. Opisując rozmaite schematy eksperymentalne, staraliśmy się pokazać, jak jedne z nich pozwalają radzić sobie z trudnościami, które były nie do pokonania na gruncie innych. Dyskutując problemy eksperymentu jako sytuacji społecznej, usiłowaliśmy pokazać, że ujawnienie uprzednio niedostrzeżonych interakcji społecznych w trakcie eksperymentu jest ważnym pierwszym krokiem do ich kontroli i pokazać próbowaliśmy różne procedury do tego służące. Przedstawiając eksperyment społeczny, opisywaliśmy pewien istotny sposób pokonywania tych trudności eksperymentu, które wynikają z małej władzy badaczy nad interesującymi ich fragmentami życia społecznego. Gdy wszystkie te sposoby nie wystarczają, to pozostaje jeszcze jeden — eksperyment zastępczy na modelach.

Eksperyment zastępczy na modelach polega na zastępowaniu niedostępnych dla manipulacji zjawisk ich modelami, badaniu modeli i przenoszeniu uzyskanych wyników z modeli na odwzorowane przez nie oryginały. Eksperyment zastępczy na modelach jest w pewnym sensie przechytrzeniem Natury: pozwala respektować dyktowane przez nią ograniczenia eksperymentu a równocześnie uzyskiwać to, co dobrego eksperyment daje. Pozostaje trzeci rodzaj reakcji na przeszkodę — znalezienie innych dróg do celu. Realizacją tej możliwości jest rozwój badań nieeksperymentalnych, przez grzeczność zwanych quasi-eksperymentalnymi. W badaniach tych zjawiska wywołuje nie badacz, lecz inne siły — praktyk czy Historia. Sam badacz jest tylko ich obserwatorem lub interpretatorem. Zamiast manipulować zjawiskami, dobiera on je, zestawia i analizuje tak, by nałożyć na nie schematy dowo-

du przyczynowego. Badanie sytuacji wyjątkowych, badania porównawcze, badania zmian społecznych o wyraźnych początkach i przebiegu, ocena wpływu reform społecznych, korelacyjna analiza wielozmiennowa danych sondażowych są przykładami takich badań. Różnica między badaniami nieeksperymentalnymi a eksperymentem nie polega na logice dowodu przyczynowego, gdyż ta jest wspólna dla wszystkich badań kontrolowanych, lecz na jego sile. Siła dowodu przyczynowego w badaniach nieeksperymentalnych jest *ceteris paribus* mniejsza niż w badaniach eksperymentalnych, bowiem choć zwykle łatwo jest ustalić korelację między pewnymi zjawiskami, to często trudno jest określić ich następstwo czasowe a jeszcze trudniej wykluczyć możliwość wyjaśnienia obserwowanych zmian innymi niż hipotetyczne czynnikami.

Metodami badań nieeksperymentalnych zajmować się nie będziemy. Ich dokładniejszy rozbiór byłby w tym miejscu i niemożliwy i niepotrzebny. Niemożliwy, gdyż określenie „badania nieeksperymentalne" jest tak pojemne, że obejmuje całość zabiegów służących do ustalania związków przyczynowych, pomniejszoną — rzecz jasna — o same eksperymenty i próba opisania ich w zakresie choćby elementarnym, wymagała pracy o objętości równej niniejszej[185]. Niepotrzebny, gdyż przedmiotem naszego zainteresowania są eksperymenty w sensie ścisłym, a nie szeroko rozumiane badania kontrolowane.

W niniejszej, ostatniej części pracy zajmiemy się właśnie analizą eksperymentów zastępczych na modelach. Podamy najpierw ogólnometodologiczną charakterystykę modelu a potem wykorzystamy ją do rozbioru dwu najczęściej spotykanych w badaniach społecznych odmian eksperymentów modelowych. W jednej z nich manipuluje się małymi grupami z myślą poznania innych grup a zwłaszcza wielkich struktur społecznych a w drugiej eksperymentuje się na zwierzętach po to, by poznać ludzi.

11. POJĘCIE MODELU I EKSPERYMENTU MODELOWEGO

Wieloznaczność pojęcia modelu jest zbyt dobrze znana, by o niej przekonywać. Zamiast więc przytaczać istniejące sposoby rozumienia tego pojęcia, z których wiele jest równoprawnych i uświęconych obyczajem językowym, obierzemy od razu pewną ich odmianę. Decyzja taka jest arbitralna z językowego punktu widzenia, lecz jest uzasadniona ze względów praktycznych. Model dla nas to ani normatywny wzorzec, ani układ odwzorowany, lecz pewna cenna ze względów metodologicznych postać samego odwzorowania.

Modelem będziemy nazywać za Sztompką „układ, który stanowi środek pośredniego poznania drugiego układu dzięki temu, że jest wobec niego analogiczny pod istotnym względem a równocześnie bardziej dostępny poznawczo"[186]. W definicji tej nie przesądza się, czym mają być przedmioty będące modelami — żąda się tylko by dobrze pełniły funkcje zastępcze. Modelami mogą być przedmioty fizykalne — celowo stworzone lub tylko obrane. Mogą być nimi i przedmioty skonstruowane myślowo — upraszczające rzeczywistość układy twierdzeń, w szczególnym wypadku —' sformalizowanych.

Walor przyjętej definicji tkwi w jej dużej dla nas użyteczności. Dostarcza ona wygodnej ramy dla ujęcia zagadnienia eksperymentu modelowego, gdyż niedostępność poznawcza oryginału, analogia między nim a modelem oraz ekstrapolacja wyników badania modelu na oryginał będą podstawowymi punktami naszej analizy eksperymentów modelowych.

*

[185] A. Sułek (oprac.), Logika analizy socjologicznej, Wyd. UW, w druku.
[186] P. Sztompka, O pojęciu modelu w socjologii, „Studia Socjologiczne" 1968, nr 1.

* *

Potrzeba korzystania z modelu powstaje wtedy, gdy chcemy zdobyć wiedzę o jakimś układzie, a ten nie jest z jakichś powodów dostępny badaniu. Przeszkodą mogą być własności skalarne układu (zbyt duże lub zbyt małe rozmiary przedmiotu, zbyt krótki lub zbyt długi czas trwania procesu), zbyt duży stopień jego złożoności, trudności techniczne, zakazy normatywne dotyczące badania itd. Wszystkie te cechy są relatywizowane do możliwości badacza. Trzeba wówczas znaleźć lub stworzyć inny układ, bardziej dostępny, którym można będzie zastąpić niedostępny oryginał, aby z jego badania uzyskać potrzebne informacje o oryginale. Do tego jednak ów substytut musi spełniać pewne warunki podobieństwa do oryginału. Przyjęta tu definicja nie odwołuje się do cech absolutnych układu kandydującego do roli modelu, ale do relacji między nim a oryginałem i do celów badania. W ten sposób każdy układ podobny do innego układu pod dobrze określonym względem może być dlań modelem a o tym, czy nim jest czy nie jest, rozstrzyga jego rola w badaniu. Stąd np. modelem mózgu może być według Turinga odpowiednio ukształtowany garnek zimnej owsianki — podobieństwo to w zupełności wystarcza do badania przechodzenia przez mózg fal uderzeniowych bez potrzeby bezpośredniego kontaktu z żywym mózgiem, gdyż mózg ma taką właśnie konsystencję[187]. Nie wystarcza ono natomiast do badania nawet najprostszych aspektów neurofizjologii mózgu ludzkiego, do czego z kolei posłużyć się można jako modelem, mózgiem zwierzęcym.

Znakomity ten przykład zwraca uwagę na kilka właściwości modeli. Po pierwsze: Jeden układ może mieć wiele modeli a każdy z nich może być pożyteczny pod pewnym ważnym dla badania względem; można o nich powiedzieć, że odwzorowują one poszczególne aspekty oryginału. Po drugie: Wszystkie modele pewnego układu można uporządkować ze względu na stopień podobieństwa do oryginału mierzony ilością cech oryginału odwzorowanych w modelu. Po trzecie: Model wierny wobec oryginału pod wszystkimi względami musiałby być oryginałem samym; jest to wniosek z Leibniza zasady identyczności. Identyczność w sensie leibnizjańskim stanowi górną granicę podobieństwa, wcale nie tak atrakcyjną zważywszy na pragmatyczne wymagania stawiane modelom: te są cenne nie tylko przez swe podobieństwo, ale i niepodobieństwo do od wzorowanych oryginałów.

Bardziej systematyczne ujęcie stosunków między modelem a oryginałem, uwzględnić musi dwie sytuacje. W sytuacji pierwszej oba układy posiadają pewne wspólne cechy. W sytuacji drugiej, między cechami tych układów istnieje tylko odpowiedniość: wspólne są nie cechy, lecz formalna struktura teorii obu układów. Teorie te są empirycznymi interpretacjami tego samego formalizmu w dwu dziedzinach — w modelu i oryginale. Pierwszy rodzaj analogii nazywa się analogią rzeczową czyli homologią a drugi nazywa się analogią formalną. W niniejszej pracy zajmować się będziemy w zasadzie tylko modelami opartymi na analogii rzeczowej. Analogie formalne w naukach społecznych są zbyt wątłe, by móc na nich polegać.

Model homologiczny posiada trzy klasy cech wyróżnionych wedle ich stosunku do cech odwzorowanego oryginału: cechy specyficzne — właściwe tylko modelowi, cechy analogiczne — wspólne obu układom i cechy nowo stwierdzone, o których jeszcze nie wiadomo, czy ma je i oryginał. Swe cechy specyficzne posiada również i oryginał. Celem badania modelowego jest znalezienie nowych cech modelu i przypisanie ich oryginałowi[188].

Jeśli model ma być narzędziem ekstrapolacji licznych i trafnych, to między wyróżnionymi klasami cech muszą zachodzić pewne stosunki. Wartości modelu nie da się ocenić przez proste porównanie liczby cech wspólnych i różnych między modelem a oryginałem. Wprawdzie wnioskowanie przez analogię — logiczna podstawa ekstrapolacji polega na tym, że z tego, iż dwa przedmioty podobne są pod

[187] Przykład wzięty z pracy W. R. Ashby'ego, Symulacja mózgu, w: H. Borko (red.), Maszyny cyfrowe w badaniach naukowych, Warszawa 1969.

[188] Zob. M. B. Hesse, Models and analogies in science, London 1963.

pewną liczbą względów a jeden z nich posiada ponadto jakąś cechę nie objętą dotąd stwierdzonym podobieństwem, wnosimy, że i drugi przedmiot cechę tę posiada, ale pewność tego wniosku nie zależy po prostu od liczby cech objętych dotąd stwierdzonym podobieństwem. Bez wątpienia, rzecz jest bardziej złożona.

Najogólniej można powiedzieć, że trafność ekstrapolacji z modelu na oryginał jest wyższa wtedy, gdy analogia między nimi jest głęboka a nie powierzchowna, istotna a nie przypadkowa. Wyrażając to porządniej, chodzi o to, by przenoszone cechy modelu posiadały swoje przyczyny lub przynajmniej silne korelaty w obrębie cech analogicznych a nie specyficznych, a zarazem, by wśród specyficznych cech oryginału nie było rozmaitego rodzaju modyfikatorów związków łączących cechy ekstrapolowane z cechami analogicznymi.

Widać więc, że poprawna ekstrapolacja z modelu na oryginał wymaga nie tylko wyraźnego uświadomienia sobie odpowiednich podobieństw między nimi, ale i różnic. Te drugie wymagają nawet o wiele większej uwagi niż pierwsze, bo chronią przed wnioskowaniami bezpodstawnymi. Różnice te można pomijać tylko wtedy, gdy są one w świetle naszej wiedzy nieistotne względem cech ekstrapolowanych. Warto wymienić przynajmniej dwa rodzaje różnic, które muszą być szczególnie troskliwie analizowane, jeśli chce się uniknąć wnioskowań bezzasadnych. Pierwszym są te, ze względu na które wybraliśmy model zamiast oryginału — wybór lub budowa modelu różnego od oryginału pod istotnym pragmatycznie względem wymaga założenia, że wzgląd ten nie jest istotny przyczynowo w stosunku do innych interesujących nas aspektów. Drugim są różnice w stopniu komplikacji między modelem a oryginałem. Jeśli model wybiera się dla jego prostoty, to trzeba liczyć się z cechami emergencyjnymi w oryginale, których pojawienie się może w poważny sposób modyfikować rzutowane zależności. Druga sytuacja jest szczególnym przypadkiem pierwszej, niemniej jednak ze względu na jej doniosłość warto było ją wyróżnić.

Trafność wnioskowania z modelu o oryginale jest większa także wtedy, gdy liczba kroków w takimwnioskowaniu jest mała. Gdy z istnienia cech wspólnych między nimi i nieistotności cech różnych wnosi się wprost o wspólnocie cech predykowanych, to zasadność wnioskowania jest wyższa niż wtedy, gdy wnioskowanie ma charakter pośredni i wymaga uwzględnienia całego łańcucha przyczynowego.

Z powyższych ustaleń wynika, że poprawne wnioskowanie z modelu zakłada nie tylko pewne rozeznanie w podobieństwach i różnicach między modelem a tym co modelowane, lecz także odpowiednie ogólne zależności empiryczne. Wiążą one cechy analogiczne ze stwierdzonymi w modelu cechami ekstrapolowanymi oraz orzekają o braku związku między tymi ostatnimi a cechami specyficznymi modelu i oryginału — stąd Black nazywa je uogólnieniami łączącymi[189].

Czasami zależności te mają charakter uogólnień sprawozdawczych, będących podsumowaniem wyników badania sytuacji, w których występowała korelacja cech analogicznych z cechami ekstrapolowanymi mimo istnienia różnic między modelem a oryginałem. Z pewnym ryzykiem błędu, zwykle towarzyszącym przewidywaniom opartym na uogólnieniach niskiego szczebla, można wówczas przypuszczać, że korelacja ta powtórzy się dla jeszcze jednej cechy. Ryzyko to jest mniejsze, gdy dysponuje się stosunkowo dobrze rozwiniętą teorią tej dziedziny zjawisk, do której należy zarówno oryginał, jak i jego model. Jeśli przypomnieć, że teoria to coś więcej niż system twierdzeń sprawozdawczych a twierdzenia jej wykraczają poza dotąd zebrane dane empiryczne i posiadają moc predyktywną, to korzyści jakie ona daje w dyskutowanej sytuacji nie wymagają objaśnienia. Zauważmy tylko, że chroni ona przed przyjmowaniem założeń *ad hoc* — te uzasadniają tylko przypadki, dla których je wprowadzono, są więc nie tylko bezwartościowe, ale i niebezpieczne.

Jasne się więc staje, że poprawność wnioskowania z modeli o oryginałach zależy nade wszystko od precyzji sformułowania i stopnia potwierdzenia zakładanych

[189] M. Black, Critical thinking, New York 1952, s. 319 - 23.

przez nie ustaleń i praw empirycznych, nie zaś od pomysłowości postępowania logicznego. Wydaje się, że o warunkach poprawności takiego wnioskowania powiedzieć można niewiele więcej. Podręczniki logiki w tym względzie albo milczą, albo każą zadowolić się mniej więcej takimi właśnie twierdzeniami. Podejmowane próby określenia logicznych warunków wnioskowania z analogii nie są ani częste ani zbyt owocne: prowadzą do wniosków z logicznego punktu widzenia mało interesujących, za to dobitnie wskazują na konieczność każdorazowego uwzględniania charakteru wchodzącej w grę sytuacji empirycznej, np. wagi podobieństw w stosunku do różnic[190]. Jeśli wziąć nadto pod uwagę, że sama relacja analogii jest zbyt słaba na to, by jej teoria była nietrywialna[191], to logiczna teoria wnioskowań z analogii nie wydaje się możliwa — pozostają one rzeczą wiedzy, pomysłowości i odwagi badaczy.

Gdyby stwierdzenia określające zakres cech analogicznych i specyficznych oraz związki z nimi nowo stwierdzonych cech modelu były i pewne i precyzyjne, i liczne, to ekstrapolacje z modeli na układy modelowane miałyby wysoki stopień trafności. Warunki te, oczywiście, nie są spełnione; „oczywiście", gdyż możliwe byłoby to tylko w sytuacji posiadania bardzo rozwiniętej wiedzy diagnostycznej i teoretycznej tak o modelu, jak i oryginale, a wtedy i samo badanie modelowe nie byłoby potrzebne. Wobec powyższego wnioskowanie modelowe ma charakter tylko hipotetyczny. Wyniki badania modelu powinny być, jeśli to możliwe, potwierdzone przez bezpośrednie obserwacje oryginału. Nigdy bowiem nie można wykluczyć tego, że specyficzne cechy oryginału zlikwidują bądź zmodyfikują zależności nań rzutowane. A tak naprawdę, to badania modelowe są najbardziej przydatne tam, gdzie mało o oryginale wiemy, gdzie zatem i ryzyko wnioskowania jest duże.

Wyniki bezpośredniego badania oryginału, jeśli takie są możliwe, są testami nie tylko dla ekstrapolacji z modelu, ale i dla przyjmowanych założeń modelowych, tj. ustaleń cech analogicznych i specyficznych oraz uogólnień łączących. Moment ten jest ważny dlatego, że założeń tych na ogół się nie podaje: wnioskowania z modelu są zatem entymematami. Przemilczanie przesłanek w tych entymematach nie bierze się z ich zakładanej oczywistości, lecz z trudności sformułowania ich w sposób precyzyjny a zarazem zgodny ze wszystkimi dostępnymi świadectwami. Przesłanki te są więc nie tylko hipotetyczne, ale i intuicyjne. Ich ujawnianie jest pracą raczej metodologów niż samych badaczy.

*

* *

Ogólnometodologiczna charakterystyka modelu i badania modelowego stosuje się także do modeli eksperymentalnych i do eksperymentów modelowych w naukach społecznych. Co więcej, stosuje się ze szczególną mocą. Eksperyment bezpośredni, będąc metodą wymagającą aktywnego i głębokiego kontaktu badacza z rzeczywistością, napotyka na szczególnie duży opór rozmaitych jej fragmentów. Eksperyment w naukach społecznych trafia na wiele dodatkowych barier nie znanych przyrodoznawstwu, uwarunkowanych nie koniecznymi cechami tej metody, lecz naturą badanych zjawisk. Sporządzony w poprzedniej części niniejszej pracy wykaz i rozbiór trudności bezpośredniego eksperymentowania na zjawiskach społecznych stanowi równocześnie wskazanie powodów, dla których nauki społeczne sięgają po eksperymenty zastępcze — na modelach.

[190] Zob. np. M. Hesse, Analogy and confirmation theory, „Philosophy of Science" 1964 vol. 31, nr 4.
[191] Zob. M. Bunge, Method, model and matter, Dordrecht 1973, s. 114 - 130.

12. MAŁA GRUPA JAKO MODEL EKSPERYMENTALNY

W obrębie rzeczywistości społecznej wyróżnia się dwa poziomy organizacji: poziom małych grup i wielkich zbiorowości, mikrostruktur i makrostruktur. Stanowią one zarazem poziomy analizy dla dwóch działów socjologii: mikro- i makrosocjologii. Choć kryteria tych zróżnicowań podać niełatwo, a i przejścia między poziomami są płynne, to jednak istnieją teorie zgodnie zaliczane do mikrosocjologii i teorie równie zgodnie zaliczane do makrosocjologii. Z ich miejscem w klasyfikacji wiążą się doniosłe charakterystyki metodologiczne. W tym miejscu szczególnie ważny jest stopień manipulowalności zjawisk, których dotyczą, określający możliwości eksperymentowania. Możliwości te są bardzo duże na poziomie małych grup a żadne na poziomie wielkich struktur. Jest to jedna z istotnych przyczyn dysproporcji rozwojowych między mikrosocjologią i makrosocjologią: teorie makrosocjologiczne rzadko mogą pochwalić się tak wysokim stopniem uzasadnienia i systematyzacji, jaki mają teorie mikrosocjologiczne.

Pojawia się pytanie o wzajemny stosunek tych dwóch działów socjologii tj. o przekładalność ich terminów i relacje logiczne między ich twierdzeniami. Jest to bardzo istotne pytanie. Gdyby bowiem można było przełożyć niektóre twierdzenia makrosocjologii na twierdzenia mikrosocjologii i odwrotnie, to pozycja socjologii wielkich struktur poprawiłaby się znacznie. Można byłoby zastąpić zbiorowości niedostępne manipulacji eksperymentalnej ich podatnymi na takie manipulacje modelami i w ten sposób, okrężną drogą, można byłoby badać za pomocą eksperymentu niektóre zagadnienia makrosocjologii. Owe hipotetyczne możliwości są w pewnym zakresie realizowane.

12.1. ŁATWOŚĆ EKSPERYMENTOWANIA NA MAŁYCH GRUPACH

Eksperymenty na małych grupach ujętych jako modele wielkich struktur pozwalają przynajmniej częściowo pokonać niektóre z omawianych wyżej barier eksperymentowania w socjologii.

1. Zmiana poziomu analizy powoduje gwałtowne osłabienie technicznych barier eksperymentowania. Możliwościami manipulacji zmiennymi makrospołecznymi dysponują tylko politycy i zorganizowane ruchy masowe, a i oni zresztą doświadczają sztywności świata społecznego. Dokonywane przez nich zmiany społeczne mają cele praktyczne, a nie poznawcze. Zmiana skali badanych zjawisk: zejście z poziomu wielkich struktur na poziom małych grup, zwłaszcza tych tworzonych przez badacza, niezmiernie rozszerza możliwości manipulacji eksperymentalnej.

Dokonana przez Malewskiego rekonstrukcja teorii socjologicznej Marksa[192] polegała m. in. na sformułowaniu jej twierdzeń w sposób czyniący je sprawdzalnymi nie tylko przez analizy historyczno-porównawcze, ale i przez badania eksperymentalne. Dla twierdzenia głoszącego, że „gdy zachodzi konflikt między interesami a uznawaną ideologią, to w skali masowej ideologia ulega modyfikacji polegającej na przystosowaniu się do interesów" możliwy jest już test eksperymentalny. W terenowym eksperymencie Thibauta systematycznie acz niesprawiedliwie nagradzano jednych a dyskryminowano drugich, co wzbudzało werbalne ataki pokrzywdzonych. W krótkim odstępie czasu uprzywilejowani unikali kontaktów z pokrzywdzonymi, gdyż uznawali te same kryteria sprawiedliwości. „Można

[192] A. Malewski, Empiryczny sens teorii materializmu historycznego, „Studia Filozoficzne" 1957, nr 2.

przypuszczać — komentuje ten eksperyment Malewski — że gdyby nierówność w podziale nagród trwała dłużej, to członkowie grup uprzywilejowanych stworzyliby ideologię usprawiedliwiającą posiadane przez nich przywileje”[193]. Sprawdzić można byłoby to łatwo. Jasne jest, że zróżnicowanie społeczne i ideologie badacz może tworzyć jedynie w zakresie bardzo a bardzo wąskim: w sytuacjach naturalnych pod mało istotnymi względami albo tylko w laboratorium.

2. Zanika bariera aksjologiczna, szczególnie — ideologiczna. Ideologia zabrania manipulacji równoważnych zmianom społecznym, których następstwa byłyby nieodwracalne i zakłócałyby rozwój lub równowagę sankcjonowanego przez nią porządku społecznego. Tymczasem w małych grupach, zwłaszcza laboratoryjnych, skutki manipulacji są i mało znaczące, i odwracalne.

Dobrym przykładem może być tu zinterpretowany modelowo eksperyment Lewina i Lippitta nad skutkami władzy autokratycznej, liberalnej i demokratycznej w grupie dla jej sprawności i integracji[194]. Okazało się m.in., że w grupie kierowanej w sposób demokratyczny jakość produktów była wyższa, więcej było wspólnego wysiłku, krytyki rzeczowej i konstruktywnej, wytwarzało się silniejsze poczucie przynależności i odpowiedzialności za grupę. Za to w grupie rządzonej w sposób autokratyczny jakość wytworów była niższa, ale silniejsze było napięcie, co znajdowało wyraz w braku ocen obiektywnych, szukaniu kozłów ofiarnych i atakach wzajemnej wrogości, silniejszy był egocentryzm, a gdy słabł wpływ przywódcy, struktura wykazywała skłonność do rozpadu. Wydaje się, ze opis ten może służyć za charakterystykę społecznych następstw demokracji i autokracji w skali społeczeństwa globalnego. Widać tu dobrze, że eksperymentalna zmiana stylu władzy możliwa jest tylko w . małych zbiorowościach społecznych, nie we wszystkich zresztą.

Bariera etyczna mniej jest wrażliwa na zmianę skali eksperymentu, ale i tu zmiana ta daje pewne korzyści. W grupach eksperymentalnych dopuszcza się w zasadzie obniżanie pozycji społecznej czy samooceny — żąda się tylko, by były to zmiany faktycznie nieistotne i przejściowe.

3. Struktura interakcji w małych grupach jest prostsza, zatem bardziej przejrzysta niż odwzorowywana struktura oddziaływań w wielkich zbiorowościach. Stąd łatwiej tak o kontrolę, jak i łatwiej o obserwację. Mała grupa jest całością, która może być ogarnięta z perspektywy badacza, natomiast społeczeństwo jest tworem konstruowanym, a nie postrzeganym. Daje to możliwość zarówno precyzyjnego wyodrębnienia zmiennych, jak również nadzorowania całości procesów społecznych. Co więcej, małą grupę łatwo izolować od zakłóceń zewnętrznych albo dobrać do niej grupy kontrolne.

4. Modelowe zmniejszenie rozmiarów badanych zbiorowości oznacza też zwykle zmniejszenie skali czasowej interesujących badacza procesów. Procesy makrospołeczne mają zwykle charakter długofalowy. Odtworzenie ich w miniaturze pozwala obserwować ich przebieg i dostrzec nawet dalekie konsekwencje działania pewnych czynników. Twierdzenie o dostosowywaniu się ideologii do interesów, eksperyment Thibauta i pomysł jego przedłużenia podsunięty przez Malewskiego mogą służyć tu jako przykład. Dopasowanie się ideologii do interesów grupy jest zwykle procesem długotrwałym — w zaaranżowanej w tym eksperymencie sytuacji można byłoby go przyśpieszyć znacznie.

[193] A. Malewski, O zastosowaniach teorii zachowania, w: O nowy kształt nauk społecznych, Warszawa 1974, s, 366- 7.

[194] K. Lewin, R. Lippitt, Eksperymentalna metoda badań nad autokracją i demokracją, w: J. N. Kosiński Socjologia amerykańska, (oprac.), Nowy Jork [1962].

12.2. KORZYŚCI MODELOWYCH EKSPERYMENTÓW NA MAŁYCH GRUPACH

Eksperymentowanie na małych grupach w miejsce niedostępnych dla eksperymentu wielkich struktur daje następujące korzyści:

1. Eksperymenty z małymi grupami albo już eksperymentalnie sprawdzone fragmenty teorii mikrosocjologicznych mogą być źródłem hipotez makrosocjologicznych. Jest to funkcja heurystyczna. Trudno podać przykład eksperymentu z małymi grupami przeprowadzonego tylko po to, by postawić hipotezy dotyczące poziomu wielkich struktur. Łatwiej o prace, które wykorzystają jako hipotezy już sprawdzone wyniki eksperymentów; z naszego punktu widzenia różnica ta jest bez znaczenia. Przykładem pracy tego rodzaju jest rozprawa Galtunga *Smali group theory and the theory of international relations: a study in iso- morphism*[195] ciekawa nie tylko dlatego, że skala analogii jest maksymalna (mała grupa — światowy system polityczny), ale też dlatego, że jej autor wyraźnie ujawnia własną świadomość i założenia metodologiczne.

Galtung uważa międzynarodowy system polityczny za dziedzinę szczególnie nadającą się do odwzorowania przez małe grupy. Zwykle tylko niewielka liczba krajów uwikłana jest w interakcje, co uzasadnia użycie słowa „mała", a system taki posiada względnie niski poziom organizacji, co uzasadnia użycia słowa „grupa". Rozwinięte wielkie struktury społeczne nie nadają się tu na modele, gdyż ich skomplikowane sieci interakcji nie mają na poziomie międzynarodowym swych odpowiedników.

Interpretacji teorii małych grup w dziedzinie międzynarodowego systemu politycznego dokonuje autor przez zastąpienie w twierdzeniach mikrosocjologicznych terminów specyficznych dla poziomu małych grup odpowiednimi terminami definiowanymi na poziomie systemu międzynarodowego. Uzyskuje w ten sposób hipotezy dotyczące stosunków międzynarodowych. Do tego musiał ułożyć słownik przekładów. I tak, jednostce odpowiada naród, grupie — sojusz, normom nieskodyfikowanym — praktyka międzynarodowa a skodyfikowanym — prawo międzynarodowe; terminy takie jak władza, interakcja czy demokracja występują w obu dziedzinach.

W ten sposób Galtung otrzymał szereg hipotez o tworzeniu się systemów międzynarodowych, władzy w tych systemach, ich wewnętrznego funkcjonowania. Hipotezy te konfrontował z danymi niezależnymi, odnoszącymi się bezpośrednio do systemu międzynarodowego. Tak np. wyniki eksperymentów Ascha specyfikujące warunki skutecznego oddziaływania przez grupę na dewiantów odniósł on również do systemu międzynarodowego a następnie określił warunki od których zależy, czy zasady te działać będą także w nowej dziedzinie.

Galtung wskazuje także na te cechy międzynarodowego systemu politycznego, które nie mają swych odpowiedników w życiu małych grup. Do takich należy m.in. brak ruchliwości przestrzennej państw, co wybitnie ogranicza możliwości przenoszenia praw rządzących tworzeniem się małych grup i skłania do szukania dla ich osobliwości nowych modeli.

2. Wartość danych eksperymentalnych z terenu małych grup w stosunku do hipotez o wielkich strukturach może być nie tylko heurystyczna, ale i uzasadniająca. Będzie tak, jeśli hipotezy makrosocjologiczne zinterpretuje się w zakresie małych grup, albo znajdzie się hipotezy odnoszące się już do obydwu poziomów, albo zweryfikowane eksperymentalnie hipotezy mikrosocjologiczne obejmują teren wielkich struktur. Ponieważ wymogi wobec tej procedury są znacznie surowsze niż w przypadku zastosowań heurystycznych (o czym niżej), więc i sama ona jest mniej zachęcająca.

[195] W tomie tego autora: Peace and world structure, Copenhagen 1977.

Pracę tego rodzaju wykonali Lewin, Lippitt i White w odniesieniu do zagadnień behawioralnych następstw autokracji[196]. Usiłowali oni odpowiedzieć na pytanie, jakie cechy struktury i kultury grupy sprawiają, że autokracja raz wiedzie do buntu przeciw władzy, innym razem do apatycznego podporządkowania się jej, a jeszcze innym do prześladowania kozłów ofiarnych. Z ich eksperymentu wynika, że czynnikami tymi są sztywność struktury społecznej i zwyczajowe sposoby postępowania. Sztywność struktury decyduje o tym, czy napięcia generowane przez władzę autokratyczną powodują odejście z grupy względnie zmianę pozycji w niej czy też, w zależności od zwyczajowych sposobów postępowania, przeradzają się w agresję wobec kozłów ofiarnych lub apatię.

W analizie wyników autorzy stawiają otwarcie sprawę ich ogólności i możliwości ich odniesienia nie tylko do małych grup, na terenie których je uzyskano, lecz także do poziomu wielkich struktur, zjawisk makropolitycznych. Piszą: „Istotą eksperymentu jest stworzenie sytuacji wykazującej pewien wzór. To, co się dzieje zależy, ogólnie mówiąc, od tego wzoru i jest w wielkiej mierze, choć nie całkowicie, niezależne od absolutnego rozmiaru pola". Możliwość ekstrapolacji przesądzono tu pozytywnie, przy czym podano wyraźnie założenie do tego uprawniające: uznaje ono taką specyficzną cechę oryginału jak duży rozmiar za nieistotną dla danej zależności.

3. Ponieważ, jak już zaznaczyliśmy, socjologia małych grup cechuje się stosunkowo dużym stopniem systematyzacji teoretycznej, co pozwala wywodzić z twierdzeń podstawowych i założeń dodatkowych wiele twierdzeń nowych, to wykazanie analogii między małą grupą a wielkim systemem w zakresie cech istotnych, pozwala bez dodatkowych badań dochodzić do twierdzeń o nowych cechach i relacjach; ich prawdziwość w nowej dziedzinie zależna będzie od prawdziwości twierdzenia o analogii i braku istotnych modyfikatorów. Z aksjomatyzacji teorii małych grup korzystać może więc i socjologia wielkich struktur.

4. Jeśli okaże się, że prawa sformułowane w jednej dziedzinie socjologii ważne są i dla drugiej, to powstaje możliwość sformułowania praw ważnych dla systemów społecznych *tout court.* W ten sposób rodzi się szansa ograniczonej choćby integracji między mikro- i makrosocjologią, dyscyplinami na pierwszy rzut oka trudno integrowalnymi. Jest to integracja nie przez redukcję jednych praw do drugich, ale przez ustalenie prawidłowości dla obydwu dziedzin. Twierdzenia takie można zaliczać do ogólnej teorii systemów społecznych, która — jak nieco emfatycznie pisze Etzioni „zajmuje się powszechnymi własnościami wszystkich układów społecznych, od grupy dwuosobowej do społeczności światowej"[197]. Przypomnimy tu jeszcze, że do budowy takich teorii nie wystarczają badania modeli; hipotezy wzięte z nich trzeba sprawdzić w sposób niezależny na terenie zjawisk makrospołecznych.

Uwagi powyższe mogły stworzyć wrażenie, że wymienione korzyści można osiągać wyłącznie za pomocą eksperymentu. Sugerowanie tego nie było naszym celem, można je osiągnąć przy pomocy badań małych grup. Badania te, choć głównie oparte na eksperymencie, korzystają także z innych metod.

12.3. ZAŁOŻENIA EKSPERYMENTU MODELOWEGO NA MAŁYCH GRUPACH

Wnioskowania z modelu (małej grupy) o oryginale (wielkiej strukturze) przyjmować muszą, przynajmniej milcząco, jakieś założenia modelowe. Ich wymagana wartość logiczna zależna jest od funkcji metodologicznej modelu: wyższa musi

[196] K. Lewin, R. Lippitt, R. K. White, Patterns of aggressive behavior in experimentally created „social climates", „Journal of Social Psychology" 1939, vol. 10, nr 2.

[197] A. Etzioni, Toward a macrosociology, w: J. S. Coleman i in., Macrosociology: theory and research, Boston 1970, s. 107.

być w przypadku zastosowań weryfikacyjnych niż heurystycznych. Założenia te to reguły przekładu jednych terminów na drugie, umożliwiające uzyskiwanie drogą podstawiania nowych terminów i nowych twierdzeń. Możliwości takich przekładów nie sposób określić bez odwołania się do świadectw empirycznych. Ani bowiem ich korzyści nie są samoczynnym uzasadnieniem ich trafności, ani też zmiana poziomu analizy sama przez się nie czyni ich sztucznymi zabiegami. Przekłady, o których mowa nie są sprawą ćwiczeń formalnych polegających na przedefiniowywaniu pojęć w drodze analitycznej. Wymagają one wiedzy przedmiotowej. Terminy sobie odpowiadające muszą mieć znaczenie już uprzednio, i to niezależnie od siebie określone. Zaś ustalenie ich odpowiedniości jest dopiero efektem stwierdzenia na drodze empirycznej, że ich desygnaty podpadają pod te same prawa empiryczne i w tych prawach terminy owe można wzajemnie zastępować. Dotyczy to także takich terminów, które brzmią jednakowo i wydają się zdefiniowanymi nie na jednym tylko poziomie analizy socjologicznej.

Wartość prób w zakresie ustanawiania analogii w dziedzinie nowych cech badanych przedmiotów zależna jest od charakteru związku łączącego je z tymi, których analogii już dowiedziono oraz od trafności założenia, że specyficzne cechy układów odwzorowywanych są dla ich wystąpienia nieistotne. Ten drugi moment jest szczególnie ważny z uwagi na ogromne różnice między małymi grupami a wielkimi zbiorowościami w typie więzi społecznej, kontroli, uczestnictwa, komplikacji struktury itd. Owe różnice wyznaczają granice ekstrapolacji zasadnych.

Formułowanie i uzasadnianie tego typu założeń uprawomocniających jest określone charakterem sytuacji badawczej, zawsze konkretnej, toteż i czynności tego rodzaju nie są rzeczą metodologów a badaczy. Zadaniem metodologów jest natomiast określenie formalnych warunków takich przykładów i wyodrębnienie odmian tych procedur. Toteż na pytanie badacza „Czy można taki np. fakt, że mała grupa skuteczniej rozwiązuje kwestie związane z demokracją, uogólniać tak dalece, by twierdzić, że jest to dowód wyższości ustroju demokratycznego nad systemem autorytarnym dla wielkiego narodu?”[198], metodolog odpowiedziałby w sposób, który pytający uznałby zapewne za wykrętny, niemniej jednak odpowiedź ta zawierałaby formalny szkielet oczekiwanej odpowiedzi.

Powtórzmy na koniec raz jeszcze, by uniknąć nieporozumień, że tylko niektóre cechy i relacje charakterystyczne dla wielkich struktur mogą być odwzorowane i badane w małej grupie. Zakres ten jednak nie jest stały, bo rośnie wraz ze wzrostem wiedzy przedmiotowej i pomysłowości badaczy.

12.4. MAŁE GRUPY SZCZEGÓLNEGO RODZAJU JAKO MODELE EKSPERYMENTALNE

Psychologia społeczna bez określników dodatkowych jest psychologią wieku dojrzałego — jej twierdzenia odnoszą się do osobników dorosłych, ludzi o osobowościach ukształtowanych. Niemniej jednak przedmiotem badań nad małymi grupami są bardzo często grupy szkolne, kolonijne czy klubowe, złożone z dzieci. Wszystkie eksperymenty, na które powoływaliśmy się w tym rozdziale, przeprowadzono na nich właśnie. Zjawisko to ma kilka przyczyn:

„Od dorosłych eksperyment wymaga dobrej woli, powagi i sporej porcji naiwności” — powiada krótko Duverger[199]. Pierwsza potrzebna jest do wykonywania poleceń eksperymentatora, druga do traktowania sytuacji eksperymentalnej serio, trzecia do zmylenia podejrzeń badanych co do faktycznego charakteru sytu-

[198] R. Bastide, Socjologia chorób psychicznych, Warszawa 1972, s. 104; być może jest to aluzja do opisywanych wyżej badań grupy Lewina.

[199] M. Duverger, Introduction to the social sciences, London 1964, s. 250.

acji, w jakiej się znajdują lub rzeczywistego celu badania. Wszystkie te elementy mają umożliwić samo przeprowadzenie eksperymentu i uniknięcie wspomnianego wyżej efektu pseudopotwierdzenia bądź pseudoobalenia hipotez. Dziećmi i łatwiej manipulować i łatwiej wprowadzić je w błąd.

Fakt, że osobowości i zachowania dzieci są stosunkowo plastyczne ułatwia badanie powolnych procesów w formach przyśpieszonych. Instruktywna w tym punkcie jest dyskusja nad metodologiczną wartością eksperymentu Malewskiego nad zależnością pewnych zachowań postdecyzyjnych od pozytywnego względnie negatywnego obrazu osoby własnej[200]. Krytyka kwestionowała wartość milczącego założenia, że osobowości badanych dzieci podlegają pod wchodzącymi w grę względami uwarunkowaniom i procesom analogicznym jak osobowości ludzi dorosłych i wskazywała na ich odmienność, na to, że obraz własnej osoby u dzieci „wydaje się być w znacznej masie wypadków chwiejny i zmienny”[201]. Obrona podnosiła natomiast, że okoliczność ta w rozważanym kontekście nie tylko nie jest wadą, lecz przeciwnie — jest zaletą, bo pozwala śledzić w formach przyśpieszonych, zintensyfikowanych i niezakłóconych powolne zazwyczaj procesy wpływu środowiska na osobowość[202].

Przyjmowane w takich ekstrapolacjach a łatwe do wyeksplikowania założenie głosi, że zachowania dzieci podlegają pod badanymi względami tym samym prawidłowościom, co i zachowania dorosłych. Założenie to, jak pokazuje przykład dyskusji nad eksperymentem Malewskiego, bywa kwestionowane. Dyskusje ta uczyniła wyraźną i tę, trywialną zresztą okoliczność, że względy dzięki którym na dzieciach łatwiej eksperymentować niż na dorosłych, są zarazem względami, których ekstrapolacja dotyczyć nie może; w eksperymencie Malewskiego można było badać uwarunkowania procesu, ale nie jego szybkość.

Jeśli w dodatku badania eksperymentalne dokonywane na dzieciach podejmuje się z myślą rzucenia światła na funkcjonowanie większych zbiorowości lub tylko ich wyniki tak modelowo się interpretuje, to liczba założeń koniecznych do wnioskowań z modelów o oryginałach rośnie, a tym samym zgodnie z podanymi wyżej kryteriami wartości takich wnioskowań, rośnie ryzyko błędu polegającego na ekstrapolacji fałszywej.

* *

„Twierdzenia o zachowaniu się ludzi — pisał Ossowski — nawet gdy formułowane są z wielkim kwantyfikatorem, nie obejmują osobników nienormalnych, nienormalnych ze względu na cechy gatunkowe, gdy uogólnienia dotyczą gatunku, albo nienormalnych ze względu na panujące wzory kulturowe, gdy uogólnienia dotyczą środowisk kulturowych”[203]. Niemniej jednak badanie osobników nienormalnych pozwala zrozumieć lepiej zachowanie osobników normalnych. Jeśli z jakichś powodów na pierwszych łatwiej eksperymentować niż na drugich, to nasuwa się możliwość wglądu w dziedziny w inny sposób może nawet w ogóle niedostępne.

Osobnikami odchylającymi się od normy, którzy wzbudzają może najwięcej zainteresowania, są pacjenci szpitali psychiatrycznych. Zdaniem Bastide’a grupy terapeutyczne z nich złożone są niezwykle wdzięcznym polem badań eksperymentalnych. „Po pierwsze dlatego, że sztuczna grupa szpitalna jest w pewnym sensie normalna lub, jeśli ktoś woli, że w szpitalu sztuczność jest naukowo uprawomocniona. Po wtóre dlatego, że zadaniem szpitala jest leczenie chorych i że wszelkie zabiegi terapeutyczne muszą uwzględniać zarówno sferę somatyczną, jak i psychiczną. Czynności, które w klasycznej socjologii małych grup były arbitralne, stają się «prawowite» ze stanowiska naukowego w obrębie murów szpitalnych. A wreszcie działanie to przybiera coraz bardziej w ostatnich latach formę terapii

[200] A. Malewski, Pozytywny i negatywny obraz własnej osoby a proces podejmowania decyzji (1962), w: O nowy kształt nauk społecznych, Warszawa 1974.
[201] Z. Zaborowski, W sprawie artykułu Malewskiego... „Studia Socjologiczne” 1963, nr 4.
[202] H. Muszyński, Eksperyment na cenzurowanym, „Studia Socjologiczne” 1963, nr 2.
[203] S. Ossowski, Zoologia społeczna i zróżnicowanie kulturowe, w: Dzieła, t. 4, Warszawa 1967, s. 348,

grupowej, to znaczy, że wymogi leczenia skłoniły psychiatrów do organizowania «małych grup»”[204]

Rola takich eksperymentów polega nade wszystko i najpierw na doskonaleniu technik terapeutycznych i poznawaniu zasad rządzących strukturą i rozwojem dewiacji. Nie możemy jednak zgodzić się z Bastidem — gdy chce ograniczyć ją tylko do tego i powiada, że „nie można uogólniać na zdrowych tego, co wynikało z doświadczeń w grupie chorych psychicznie”[205]. Zastrzeżenia nasze są następujące:

Pierwsze, dość zresztą banalne spostrzeżenie głosi, że eksperymenty terapeutyczne dostarczają także wiedzy o etiologii chorób psychicznych, a trudno powiedzieć kogo właściwie etiologia dotyczy: zdrowych czy chorych? Pomijając trudności definicji choroby, bo nic o to tu chodzi, powiedzieć trzeba, że zdrowie i choroba to terminy oznaczające tylko punkt początkowy i końcowy, względnie odwrotnie, tego karnego procesu.

Co jednak ważniejsze, można wskazać sytuacje, w których jedyną drogą eksperymentalnego uprawomocnienia hipotez dotyczących osobników normalnych jest eksperyment terapeutyczny. Przykładem metodologicznie pięknym, choć moralnie może okrutnym niech będzie doświadczenie, które jest tematem książki Rokeacha o trzech Chrystusach z Ypsilanti[206]

Celem teoretycznym tego eksperymentu było zbadanie w dobrze kontrolowanych warunkach następstw podważenia podstawowych przekonań w systemie wyobrażeń człowieka dla funkcjonowania jego osobowości. Takim przekonaniem jest przekonanie o własnej identyczności. Obserwacje poczynione np. w obozach koncentracyjnych sugerowały, że utrata poczucia własnej identyczności pociąga za sobą rozpad całego systemu przekonań, a w sytuacji skrajnej — schizofreniczne rozszczepienie osobowości. Dla oczywistych powodów moralnych obserwacji tych sprawdzać eksperymentalnie nie było wolno.

Zatem Rokeach spróbował uczynić to na schizofrenikach, przypuszczał bowiem, że następstwem odebrania im fałszywego poczucia tożsamości będzie wywrócenie się całego systemu przekonań i powrót do zdrowia — odwrotnie niż u osób zdrowych. W tym celu postawił chorych wobec konfliktu dwóch przekonań podstawowych, że są oni tymi, za kogo się uważają tj., Chrystusami i że może istnieć tylko jeden człowiek o tej samej identyczności. Oczekiwał, że któryś z chorych uważających się za Chrystusa, stanąwszy wobec faktu istnienia innych osób żywiących to samo przekonanie odrzuci je i wróci do zdrowia, potwierdzając zarazem hipotezę teoretyczną, na której oparto próbę terapii. Eksperyment skończył się niepowodzeniem. I tak jednak widać z jego opisu, że w pewnych sytuacjach empirycznych, istotnych dla sprawdzenia hipotez psychologicznych, z założenia nie można spotkać w badaniach osobników normalnych. Co było do udowodnienia.

12.5. GRY SYMULACYJNE

W ostatnich latach do sprawdzania teorii społecznych zaczęto stosować gry symulacyjne, od dawna już używane do innych celów, takich jak np. przedstawianie działania złożonych systemów organizacyjnych czy ćwiczenie się w podejmowaniu decyzji. W grach prowadzonych w celu sprawdzania teorii społecznych aranżuje się pewne sytuacje laboratoryjne, które uznaje się za modele szerszych struktur i procesów. Podstawowy problem metodologii takich badań nie polega na konstrukcji technik eksperymentalnych — te mogą być zapożyczone z badań laboratoryjnych na małych grupach, lecz na takim przeprowadzeniu gry, by z jej przebiegu i wyników można było wyciągnąć mocne wnioski na temat procesów społecznych, które dana gra modeluje.

[204] Bastide, op. cit., s. 104.
[205] Ibidem, s. 105.
[206] M. Rokeach, The three Christs of Ypsilanti, New York 1964; w Ypsilanti znajdował się szpital psychiatryczny.

Metodologia gier symulacyjnych podejmowanych dla sprawdzenia teorii społecznych nie jest jeszcze — zdaniem kompetentnych autorów[207] — rozwinięta. Istnieją dopiero pierwsze jej próby. Na jednej z nich — wykonanej przez Walaszek[208] — oprzemy dalsze uwagi.

Walaszek zarzuca dotychczasowym ujęciom metodologii gier symulacyjnych, że nie uwzględniają one relacji między teorią a grą, wskutek czego nie dostarczają podstaw do porównań między wynikami gry a implikacjami teorii. Ujęcie zaproponowane przez autorkę ma być wolne od tej wady.

Układ odniesienia pozwalający analizować gry symulacyjne składa się z pewnych elementów i z pewnych relacji między nimi. Elementy te to: badana rzeczywistość, teoria (-e) rzeczywistości, projekt gry symulacyjnej, wyniki gry i predykcje z teorii. Relacje omówimy szczegółowiej, gdyż one są tu najważniejsze.

Przede wszystkim istnieje relacja między badaną rzeczywistością a jej teorią (-ami). Teoria jest produktem abstrakcji. Dalej, istnieje relacja między teorią (-ami) a przewidywaniami dotyczącymi określonej sytuacji empirycznej. Przewidywania wyprowadza się w drodze dedukcji logicznej.

Najważniejsze dla zrozumienia czym jest gra są relacje miedzy teorią a projektem gry oraz między rzeczywistością a tymże projektem. Projekt gry jest przez te dwa elementy wyznaczony. Z jednej strony musi on operacjonalizować w języku działań badanych założenia teorii, a z drugiej musi odwzorować istotne cechy rzeczywistości — zoperacjonalizowanym pojęciom teorii przypisuje się więc pewne wartości wzięte z rzeczywistości.

W grze zaprojektowanej przez Walaszek a nazwanej „Establishment" teorią była teoria działania w sformułowaniu Colemana, a konkretną symulowaną dziedziną rzeczywistości struktura partii politycznej. Istotne dla tej teorii pojęcie aktora, w grze znalazło operacyjny odpowiednik w osobach badanych symulujących zachowanie członków „establishmentu" — ciała podejmującego decyzje oraz w osobach badanych symulujących zachowanie przedstawicieli społeczeństwa zainteresowanego w tych decyzjach. Odpowiednikiem innego pojęcia — zdarzenia były „kwestie", np. rozwój szkolnictwa lub zmniejszenie wydatków na obronę oraz dobra prywatne, np. posiadanie willi. Kontrola polegała w grze na możliwości głosowania za lub przeciw „kwestiom" lub na posiadaniu i nieposiadaniu dóbr. Kontrola była zróżnicowana dla poszczególnych grających. Grającym zadano pewne cele do osiągnięcia — cele te polegały na zdobyciu kontroli. Mogli oni dla ich realizacji wchodzić w kontakty z innymi, wymieniać kontrolę i zawiązywać koalicje. Operacyjnie przybierało to postać wymiany żetonów. Ustalono jednak pewne reguły gry, które miały sprawić, że będzie ona symulacją procesu politycznego w hierarchicznie zorganizowanej strukturze. Tak np. samo istnienie hierarchii przełożono na nierówny podział zasobów', przywileje związane z pozycją — na ułatwienie pewnym grającym dostępu do dóbr prywatnych kosztem innych grających itd.

Następna relacja zachodzi między projektem gry a jej wynikami. Zaprojektowaną grę poddaje się eksperymentowaniu laboratoryjnemu i rejestruje się działania i wyniki grających. Grający trzymając się pewnych reguł, mają zarazem pozostawioną dużą swobodę decyzji. Te zależą od ich motywacji, pomysłowości, stopnia identyfikacji z rolą itd. — jest to spontaniczny aspekt gry. W grze „Establishment" w wyniku eksperymentowania uzyskano pewien empiryczny rozkład kontroli końcowej sprawowanej przez grających a także empiryczne oszacowanie wartości zdarzeń i władzy grających. Zachowania grających symulowały proces polityczny w skali makrospołecznej.

Ostatnia relacja w grze to stosunek wyników gry od predykcji wyprowadzonych z teorii. Zgodność wyników z predykcjami stanowi częściowe potwierdzenie teorii a gdy teorii jest więcej, to ta z nich jest najlepsza, której predykcje są najbliższe wynikom gry. Pamiętając o tym, że teoria opisuje rzeczywistość, zyskujemy

[207] Zob. J. S. Coleman, Recent developments in American sociological methods. „Polish Sociological Bulletin" 1974, nr 2.

[208] Z. Walaszek, On methodology of gaming: testing theories in the absence of field research. „Polish Sociological Bulletin" 1975, nr 1 - 2.

tym samym wiedzę o owej rzeczywistości, choć z nią samą mieliśmy kontakt tylko intelektualny — eksperymentowaliśmy na jej modelu.

13. EKSPERYMENTY MODELOWE NA ZWIERZĘTACH

Modelami przydatnymi do badania wielu aspektów zachowania ludzkiego są zwierzęta. Zamiast eksperymentować na ludziach, eksperymentuje się na łatwiejszych do manipulowania zwierzętach i przenosi się — w granicach analogii między modelami a oryginałami — wyniki tych doświadczeń na ludzi. W rozdziale tym przedstawimy kolejno: zakres modelowego eksperymentowania na zwierzętach; powody, dla których na zwierzętach łatwiej eksperymentować niż na ludziach; korzyści, jakie nauki o zachowaniu ludzkim mogą wynieść z eksperymentów na zwierzętach; założenia, jakie muszą być spełnione, aby korzyści te można było osiągać oraz ograniczenia i kierunki krytyki takich eksperymentów.

13.1. ZAKRES MODELOWEGO EKSPERYMENTOWANIA NA ZWIERZĘTACH

Na modelach - zwierzęcych bada się wiele ważnych wymiarów indywidualnego i społecznego zachowania człowieka: ewolucję i funkcjonowanie układu nerwowego, spostrzeganie, motywacje i emocje, procesy poznawcze, uczenie się i socjalizację, konflikt, przystosowanie i zdrowie psychiczne a nawet najprostsze więzi społeczne. Rolę modelu pełnią różni przedstawiciele królestwa zwierząt: małpa, pies, owca, żółw, szczur, gołąb, kura, ryba — że wymienię tylko najczęściej używane. Modelowe badania zwierząt prowadzi się zarówno w warunkach laboratoryjnych jak i naturalnych. Zwykle są to badania eksperymentalne[209].

Eksperymenty na zwierzętach tkwią u początków pewnych wpływowych kierunków w naukach o zachowaniu ludzkim. Badania Pawłowa nad powstawaniem odruchów warunkowych u psów dostarczyły danych wyjściowych dla teorii klasycznego uczenia się, badania Kohlera nad rozwiązywaniem problemów przez szympansy to początki psychologii postaci. Niektóre szkoły nauk o zachowaniu programowo badały zwierzęta w celu zrozumienia zachowania ludzi. Behawioryści, upatrujący w świadomości flogiston nauk społecznych i odrzucający introspekcję na rzecz pomiaru zachowań zewnętrznych, znaleźli w zwierzętach obiekty, których badanie pozwalało im — zauważa Ossowski — nie tylko realizować swe wzorce metodologiczne, lecz także ćwiczyć się i przygotowywać do badania ludzi — takimi samymi metodami[210].

Co więcej, różnice w treści między teoriami zachowania człowieka mogą być w części spowodowane różnicami między gatunkami zwierząt eksperymentalnych, najczęściej używanymi przez teoretyków. Faktem jest, że sławni eksperymentatorzy mieli swoje ulubione gatunki eksperymentalne: Pawłow — psa, Kohler i Harlow — małpę, Skinner — szczura i gołębia. Ich wybór był zapewne podyktowany przekonaniem, że procesy istotne dla zachowania ludzkiego występują u tych zwierząt szczególnie wyraźnie. Ale możliwy jest też i wpływ odwrotny: raz przyjęty model określał perspektywę widzenia a w konsekwencji i treść teorii. Tak np. doświadczenia z małpami sugerują bardziej złożone mechanizmy regulacji zachowania ludzkiego niż doświadczenia ze szczurami.

[209] Zob. Środowisko a życie psychiczne. Wybór artykułów z „Scientific American", Warszawa 1969.

[210] S. Ossowski, Zoologia społeczna i zróżnicowanie kulturowe, w: Dzieła, t. 4, Warszawa 1967, s. 344 - 5.

Używanie zwierząt jako modeli w badaniach naukowych, nie tylko behawioralnych, ale także fizjologicznych i farmakologicznych jest tak pospolite, że eksperymentowanie bezpośrednio na ludziach bywa krytykowane jako czyniące z ludzi „zwierzęta laboratoryjne". Tymczasem to właśnie zwierzęta zastępują w laboratoriach ludzi, a nie odwrotnie.

13.2. ŁATWOŚĆ EKSPERYMENTOWANIA NA ZWIERZĘTACH

W punkcie tym odpowiemy na pytanie, dlaczego na zwierzętach łatwiej eksperymentować niż na ludziach. Odpowiedź ta, połączona z pokazaniem pożytków z doświadczeń zastępczych na zwierzętach będzie wyjaśnieniem korzystania z nich przez nauki o zachowaniu ludzkim.

1. Interesujące badaczy mechanizmy regulujące zachowania zwierząt są prostsze niż ich ludzkie odpowiedniki. Działań ich nie modyfikują doniosłe w sytuacji człowieka wyższe czynności psychiczne związane z istnieniem języka i dziedzictwa kulturowego. Toteż procesy występujące u ludzi w postaci rozwiniętej, złożonej i ukrytej, u niższych biologicznie zwierząt widoczne są w formach elementarnych, prostych i jawnych. Jeśli zważyć, że istotą eksperymentu jest upraszczanie złożonych sytuacji naturalnych tak, aby można było badać zjawiska i procesy w stanie czystym, to w przypadku eksperymentów zastępczych na zwierzętach cel ten osiąga się przez samą tylko zamianę przedmiotu badania.

2. Używanie zwierząt zamiast ludzi pozwala badaczowi pokonać barierę technicznej niemanipulowalności przedmiotami badania. Społeczna rola badacza i środki jakimi dysponuje umożliwiają mu wszystkie manipulacje technicznie wykonalne, jego władza nad obiektami badania jest całkowita, badacz jest dosłownie panem ich życia i śmierci.

Niezmiernie rozszerza się wachlarz wykonalnych stymulacji eksperymentalnych, polegających tak na wprowadzaniu czynników zewnętrznych, jak i na hamowaniu działania czynników naturalnych. Tak więc badacz może zarówno poddawać zwierzę działaniu wymyślonego bodźca, choćby wstrząsu elektrycznego, jak też pozbawiać je możliwości zaspokojenia potrzeb pierwotnych, np. potrzeby zaspokojenia głodu. Wybitnie wzrastają też możliwości kontroli eksperymentalnej: łatwo można izolować stworzone sytuacje od zakłóceń zewnętrznych i izolację tę długo utrzymywać, natomiast utrzymywanie ludzi w wyjątkowych, izolowanycn sytuacjach nie jest — rzecz jasna — dla badacza wykonalne. Pomnożenie możliwości stymulacji przez możliwości kontroli pozwala eksperymentatorowi dowolnie kształtować i nadzorować cały przebieg życia zwierząt a nawet tworzyć — w granicach zakreślonych stanem wiedzy — pożądane kombinacje genetyczne.

Jeśli badacz chce np. ustalić udział dziedziczności i środowiska w kształtowaniu zdolności widzenia u człowieka, zbadać jak człowiek uczy się posługiwać wzrokiem, w jaki sposób nieużywanie wzroku hamuje rozwój jego zdolności widzenia itp., to może trzymać zwierzę w całkowitej ciemności od urodzenia aż do czasu badania. Ze sposobności tej skorzystał z powodzeniem Riesen[211]. Obserwacja ludzi-pacjentów po usunięciu wrodzonej zaćmy nie dała mu jednoznacznych odpowiedzi na owe pytanie z powodu m.in. ich niepełnej ślepoty w okresie przedoperacyjnym; posłużył się więc osobnikami pozbawionymi w sposób sztuczny i skuteczny możliwości widzenia. Badacze nie posiadają władzy potrzebnej do wykonywania takiego zabiegu na ludziach (pomijamy w tym miejscu zakazy etyczne), toteż Riesen użył szympansów, podobnych do ludzi pod wchodzącymi w grę względami.

Niektóre gatunki zwierząt posiadają jeszcze dodatkowe przymioty czyniące z nich szczególnie cenne modele eksperymentalnych badań nad różnymi dziedzinami zachowania ludzkiego. Tak np. szczur biały jest tani i łatwy do hodowli, zajmuje mało miejsca w laboratorium, łatwo można go izolować od bodźców

[211] A.H. Riesen, Niedorozwój funkcji wzroku, w: Środowisko a życie psychiczne, op. cit.

zewnętrznych, debrze znosi zamknięcie i doń się przystosowuje, bez trudu podporządkowuje się manipulacji i nie popada w nerwice pod wpływem samego eksperymentowania. Zalety te sprawiły, że jest on najchętniej używanym modelem w badaniach uczenia się[212]. Natomiast emocje i nerwice lepiej jest badać na zwierzętach wrażliwych na stres — małpach, psach, owcach.

3. Radykalnie słabnie aksjologiczna bariera manipulacji. Piwne wartości krępujące badacza w eksperymentach na ludziach, np. prywatność i pozytywna samoocena badanych, w eksperymentach na zwierzętach w ogóle nie występują. Inne wartości, takie jak życie i zdrowie istot badanych, pozostają drogimi sercu badacza, jednakże stają się mniej kategoryczne niż w badaniach ludzi. Tam nie wolno mu łamać ich pod żadnym warunkiem. W przypadku zwierząt można poświęcić je dla dobra badania, jeśli jest ono dostatecznie ważne i rzeczywiście tego wymaga. Wolno mu więc pozbawić zwierzęta kontaktu z ważnymi dla ich rozwoju składnikami środowiska i umieszczać je w niezwykłych, a nawet groźnych dla nich warunkach, aby przez obserwację następstw wywołanych zmian sformułować odpowiednie twierdzenia przyczynowe.

Pozbawienie rozwijającego się człowieka związków z pewnymi fragmentami środowiska społecznego byłoby najprostszą drogą do określenia ich znaczenia dla kształtowania się osobowości i umiejętności ludzkich. Na przeszkodzie takim doświadczeniom staje jednak dobro badanych i groźba spowodowania nieusuwalnych deformacji osobowości, bo tym próby owe często by się kończyły. Skutki, o których mowa, można natomiast określić precyzyjnie i pewnie na podstawie eksperymentów na zwierzętach. I tak np. Harlow w serii sławnych eksperymentów badał znaczenie kontaktu dziecka z matką dla kształtowania się jego uczuciowości[213]. Wśród wielu wniosków z tych badań znalazło się również twierdzenie, że oddzielenie dziecka od matki we wczesnym okresie rozwoju czyni je niezdolnym do tworzenia trwałych związków uczuciowych w przyszłości. Rzecz jasna, doświadczenie zakończone takim wnioskiem wolno było Harlowowi przeprowadzić tylko na zwierzętach. Za modele służyły mu nowo narodzone małpki, lecz wnioski z ich badania odnoszą się — jego zdaniem — także do dzieci ludzkich.

Przykłady eksperymentów zakazanych na ludziach a dozwolonych na zwierzętach można z łatwością mnożyć. Szczególnie efektownych dostarczają badania z zakresu etiologii zaburzeń psychicznych. W doświadczeniach tych wywołuje się umyślnie zakłócenia w funkcjonowaniu układu nerwowego i osobowości oraz trwałe schorzenia organiczne istot badanych. W jednym z eksperymentów tego rodzaju Brady[214] pokazał, że permanentny stan napięcia psychicznego u małp spowodowany koniecznością unikania dawkowanych im wstrząsów elektrycznych prowadzi do owrzodzenia przewodu pokarmowego, a nawet do śmierci zwierząt, jeśli okresy stresu i odpoczynku występują po sobie w pewnym określonym rytmie. Wyniki tego eksperymentu są istotne dla zrozumienia źródeł chorób zawodowych ludzi obarczonych odpowiedzialnością i pracujących w ciągłym napięciu — sam eksperyment wykonalny jest jednak tylko na zwierzętach postawionych w takiej właśnie sytuacji.

Dla ścisłości zauważyć trzeba, że groźne dla zdrowia ludzkiego i stąd prowadzone na zwierzętach eksperymenty nauk behawioralnych budzą niekiedy protesty etyczne. Same eksperymenty Harlowa były wymieniane jako wymowny przykład nagannego rozminięcia się wartości nauki z wartościami ludzkimi[215]. Wpływ tych krytyk na praktykę doświadczalną jest jednak niewielki.

4. W eksperymentach zastępczych na zwierzętach nieporównanie mniejsze niż w badaniach z ludźmi jest pozastymulacyjne, niemierzone działanie eksperymentatora na istoty badane. Aczkolwiek znane są wielce wymowne przykłady doświadczeń, w których zwierzęta reagowały nie tyle na bodziec eksperymentatora, co na jego niekontrolowane sygnały mimiczne czy głosowe (omówiony wyżej *casus* konia

[212] Zob. B. F. Skinner, The behavior of organisms, New York 1933, s. 47 - 8.
[213] H. F. Harlow, Miłość u dzieci rezusów, w: Środowisko a życie psychiczne, op. cit.
[214] J. V. Brady, Owrzodzenie przewodu pokarmowego u małp „odpowiedzialnych", ibidem.
[215] Zob. C. Roberts, The scientific conscience, New York 1967, rozdz. 6.

Sprytnego Jasia jest przykładem najbardziej wymownym), to jednak uświadamianie sobie przez zwierzęta celów badania i umyślne dostosowywanie swego zachowania do hipotezy imputowanej eksperymentatorowi lub też działanie umyślnie tę hipotezę wywracające — częste w eksperymentach z ludźmi — są tutaj z gruntu wykluczone. Wyniki uzyskiwane w eksperymentach zastępczych na zwierzętach stają się tym samym bardziej trafne.

5. Eksperymentowanie na zwierzętach — modelach ludzi nie tylko ułatwia wywoływanie procesów, które u ludzi spowodować trudno, lecz także ułatwia obserwację procesów, które na ludziach trudno obserwować. Wśród tych zaś są takie, których na ludziach wywołać nie sposób, ale są i takie, które co prawda spowodować by można, lecz mimo to trudno byłoby je nadzorować i dokładnie rejestrować. Jeśli jakieś zmiany dokonują się w organizmach i grupach ludzkich zbyt wolno lub zbyt szybko na to by je móc całościowo objąć i skrupulatnie prześledzić, to gdy można je wywołać w modelach-zwierzętach, zyskuje się niekiedy pożądane skrócenie lub wydłużenie czasu ich trwania, a więc możliwość badania za pomocą eksperymentu procesów nawet bardzo szybkich lub bardzo powolnych.

Również i z tego powodu wiele z procesów uczenia się łatwiej badać na zwierzętach niż na ludziach: zwierzęta na ogół uczą się wolniej. Z kolei pewne długofalowe procesy demograficzne można poznawać (a przynajmniej ilustrować) za pomocą nie tylko historycznych analiz rozwoju populacji ludzkich, lecz również eksperymentalnych badań prawidłowości rozwoju zbiorowości organizmów wyróżniających się dużą rozrodczością i szybką zmianą pokoleniową — organizmem spełniającym te warunki i stąd stosowanym w takich sytuacjach jest muszka owocowa.

Ciekawą ilustracją korzyści, jakie tą drogą można osiągnąć są znowu doświadczenia Harlowa nad kształtowaniem się przywiązania dziecka do matki. Aby określić czynniki za nie odpowiedzialne trzeba byłoby obserwować jego reakcje już w pierwszych tygodniach życia niemowlęcia na eksperymentalnie wywołane pojawienie się lub brak różnych elementów kontaktu z matką, takich jak karmienie piersią, kołysanie czy ciepło ciała. Tymczasem rozwój fizyczny niemowlęcia nie nadąża za jego rozwojem emocjonalnym. W pierwszych miesiącach życia, kiedy to owo przywiązanie dziecka do matki zaczyna się tworzyć, koordynacja ruchu u dziecka jest zbyt słaba, by móc jego reakcje ocenić. Kiedy zaś dziecko osiągnie wreszcie tę umiejętność, to jego uczucie przywiązania do matki jest już rozwinięte i na badanie jest już za późno. Natomiast rozwój fizyczny dzieci małp jest szybszy, ich reakcje można badać już w 10 dni po urodzeniu, co umożliwiło Harlowowi przeprowadzenie na nich eksperymentów zastępczych. Co więcej, szybsze dojrzewanie małp umożliwia obserwację następstw pozbawienia dziecka kontaktu z matką lub niektórymi jej atrybutami na jego funkcjonowanie w późniejszym okresie życia. Podobny eksperyment z ludźmi nawet gdyby był wykonalny, musiałby trwać zbyt długo jak na cierpliwość badaczy i możliwości kontroli zachowania badanych.

Dzięki tym oto okolicznościom eksperymentowanie na zwierzętach — modelach jest łatwiejsze niż eksperymentowanie na ludziach — oryginałach. Pokażemy teraz, jakie znaczenie dla nauk o zachowaniu ludzkim ma korzystanie z tej sposobności rozszerzania granic eksperymentu.

13.3. KORZYŚCI Z EKSPERYMENTOWANIA ZASTĘPCZEGO NA ZWIERZĘTACH

1. Wyniki zastępczych eksperymentów na zwierzętach są ważne jako inspiracje — bywają one hipotezami dotyczącymi zachowania ludzkiego. „Jeśli nawet takie dane muszą być zweryfikowane na materiale ludzkim, to jednak często łatwiej jest zaobserwować występowanie pewnych regularności, po uprzednim zbadaniu ich w sytuacjach prostszych i bardziej izolowanych, gdyż wówczas wiadomo już do-

kładnie czego szukać" — piszą Dollard i Miller[216]. Ich prace zawierają wiele przykładów czerpania hipotez psychologicznych z doświadczeń na zwierzętach[217]. Tę korzyść przyjęło się nazywać heurystyczną.

Wartość wyników doświadczeń na zwierzętach dla nauk o zachowaniu ludzkim wykracza jednak daleko poza pożytki heurystyczne.

2. Wyniki badań dyskutowanego rodzaju występują nader często w roli uzasadniającej względem hipotez o zachowaniu ludzkim. „W wypadkach, w których nie można przeprowadzić na materiale ludzkim eksperymentów dostatecznie ściśle kontrolowanych, przekonanie o prawdziwości wyników utwierdza się, jeżeli są one zgodne z wynikami ściślej kontrolowanych eksperymentów, przeprowadzonych z innymi ssakami"[218]. Cytowane w części poprzedniej doświadczenia Riesena i Harlowa podjęte były właśnie z taką myślą i ich autorzy nie próbowali nawet sprawdzać swych wyników na ludziach, gdyby zresztą mieli takie możliwości, to skorzystaliby z nich od razu bez uciekania się do modeli zwierzęcych.

Oprócz tak ogólnego opisu uzasadniającej roli wyników doświadczeń na zwierzętach wobec hipotez o zachowaniu ludzi, warto wyróżnić dwie ciekawe sytuacje metodologiczne, w których rola owa pełniona jest ze szczególnie dużym powodzeniem. Oto one:

2a. Wedle pewnej koncepcji metodologicznej teoria naukowa jakiejś dziedziny zjawisk jest ciągiem twierdzeń idealizacyjnych, uporządkowanych wg wagi czynników, o których twierdzenia te głoszą. Wśród owych czynników są takie, które uznaje się za podstawowe dla innych i takie, które uważa się za modyfikatory poprzednich. Odpowiednio, różne fragmenty teorii traktuje się jako opisy tej samej dziedziny, tyle, że niejednakowo dokładne. Tak więc niektóre twierdzenia dosyć dokładnie opisujące zachowanie zwierząt są zarazem przybliżonymi opisami odpowiednich zachowań ludzkich. Zdają one sprawę z procesów pod podstawowymi względami sobie podobnych a wyznaczonych przez wspólne struktury organiczne i wspólne doświadczenia, jednakże pomijają modyfikacje powodowane przez wyższe czynności psychiczne i kulturę. Natomiast rozwijanie teorii polega na uwzględnianiu owych początkowo pominiętych modyfikatorów.

Prawa teorii uczenia się niech będą tu przykładem. Homans zaczyna wykład swej teorii zachowań społecznych od prezentacji podstawowych jej praw, wywiedzionych z badań gołębia — „ulubionego zwierzęcia eksperymentalnego [jego] ulubionego psychologa eksperymentalnego B. F. Skinnera"[219].

2b. W niektórych sytuacjach metodologicznych wyniki doświadczeń na zwierzętach są czymś więcej niż nawet dobrym pierwszym przybliżeniem do praw ludzkiego zachowania. Jeśli przyjmuje się, że zachowania zwierzęce są granicznymi przypadkami przejawiania się pewnych prawidłowości ludzkich, to stwierdzenie owych regularności w świecie zwierzęcym czyni zbędnym ich poszukiwanie w świecie ludzkim. Można wtedy wypowiedzieć formułę: *a fortiori* występują one również i u ludzi. Badania ludzi określają wtedy już tylko siłę czy też modyfikacje zależności. Sama ona jest już w świetle uprzedniej wiedzy przyjęta.

Jeżeli, powiedzmy, zakłada się, że motywacje ludzkie są bardziej złożone niż zwierzęce a przy tym uznaje się pewną ich teorię, to eksperymentalne stwierdzenie, że motywacje zwierzęce są bogatsze niż zakładane przez tę teorię, prowadzi nieuchronnie do uznania tej teorii za uproszczoną, jeśli już nie fałszywą. Przykładem niechaj będą dociekania nad naturą czynników decydujących o przywiązaniu niemowlęcia do matki. Psychoanalitycy i behawioryści upatrywali je w uldze powodowanej zaspokojeniem przez matkę jego głodu. Harlow eksperymentując z małpami dowiódł, że czynnik ten jest drugorzędny nawet u zwierząt, przywiązanie zaś

[216] J. Dollard, N. E. Miller, Osobowość i psychoterapia, Warszawa 1967, s. 65.

[217] Oprócz Osobowości i psychoterapii wymienić trzeba studium Millera Psychoanalityczne przemieszczenie agresji a generalizacja bodźca i reakcji, w: A. Malewski (red), Zagadnienia psychologii społecznej, Warszawa 1962.

[218] N. E. Miller, Liberalization of basic S-R concepts, w: S. Koch (ed.), Psychology. A study of science, New York 1959, vol. 2, s. 203,

[219] C. G. Homans, Social behavior. Its elementary forms, New York 1961, s. 18.

do matki nie jest wyuczone; ale wrodzone, a czynnikiem najważniejszym w wyborze obiektu przywiązania jest potrzeba kontaktu z ciepłym i miękkim ciałem. Mógł on w konsekwencji przyjąć, że czynnik nutrytywny będzie miał jeszcze mniejsze znaczenie u ludzi, a tym samym mógł uznać poprzednie teorie za fałszywe[220].

3. Ustalenia z doświadczeń na zwierzętach pełnią także funkcje wyjaśniające względem twierdzeń o ludziach. Jakie typy wyjaśniania mogą wchodzić tutaj w grę?

Bez wątpienia można tu mówić o wyjaśnianiu w sensie psychologicznym, które polega na sprowadzeniu tego, co nieznane do tego, co znane. Takie wyjaśnienie jest jednak pojęciem pragmatycznym, bo zakres tego, co znane i tego, co nieznane musi być relatywizowany do podmiotu, któremu się wyjaśnia, jego wiedzy i możliwości przyjmowania faktów nowych. Zatem odwołanie się do praw opisujących zachowanie zwierzęce będzie psychologicznym wyjaśnieniem tylko dla kogoś, kto prawa te już sobie przyswoił, ale i wtedy wyjaśnione będzie tylko to, co ludziom i zwierzętom wspólne a to, co dla ludzi specyficzne nie znajdzie takiego wyjaśnienia. Wyjaśnienie tego typu może mieć natomiast walor dydaktyczny, gdyż eksplanans jako prostszy, jest łatwiej przyswajalny.

Można tu mówić także o wyjaśnianiu w sensie metodologicznych. Polega ono na włączeniu eksplanandum w pewien system twierdzeń. Stosunki między eksplanandum a eksplanansem nie mają tu charakteru psychologicznego jak poprzednio, lecz logiczny.

Rodzajem wyjaśniania w sensie metodologicznym jest wyjaśnianie genetyczne. Polega ono na wskazaniu wcześniejszych faz rozwojowych układu, którego działanie chcemy wyjaśnić oraz praw rządzących przejściem z jednej fazy do drugiej. W przypadku rozważanym, prawidłowości zachowania ludzkiego, w określaniu których ważną rolę odgrywają język, myślenie i kultura, byłyby wyjaśniane przez wskazanie prawidłowości zachowania się wcześniejszych ewolucyjnie zwierząt, u których czynniki te jeszcze nie działają. Zaznaczyć jednak trzeba, że w wyjaśnieniach tego rodzaju nie podaje się wymaganych przez schemat eksplanacji praw opisujących historyczne mechanizmy powstania człowieka, a zatem wyjaśnienia takie są w istocie tylko szkicami wyjaśnień w sensie Hempla[221]. Okoliczność ta nie odbiera wyjaśnieniom genetycznym ich wartości, lecz tylko ją ogranicza.

Pytaniem jest, czy można tu mówić także o najważniejszym rodzaju wyjaśniania w sensie metodologicznym, tj. o wyjaśnianiu nomologiczno-dedukcyjnym. Polegałoby ono tutaj na logicznym wywodzeniu praw o zachowaniu ludzi z — ogólniejszych od nich — praw o zachowaniu zwierząt. Z wyjaśnieniem tego rodzaju wiążą się dwa zastrzeżenia: a) wyjaśnianie dotyczyć będzie tylko tych aspektów zachowania ludzkiego, które znajdujemy w świecie zwierzęcym; b) założy się na podstawie świadectw niezależnych, że prawa dotyczące zwierząt stosują się i do ludzi. Spełnienie warunku drugiego jest jednak równoznaczne z uzasadnieniem analogicznych praw w dziedzinie zachowań ludzkich, a wówczas wyjaśniać się będzie już nimi, a poprzednie, po odegraniu swej roli heurystycznej, okażą się zbędne. Co najwyżej, stwierdzenie występowania pewnej prawidłowości w obydwu dziedzinach okaże się bodźcem do zastanawiania się nad tym, jakie to bardziej podstawowe a wspólne obydwu poziomom czynniki o tym decydują.

4. Eksperymentowanie na zwierzętach służy także usprawnianiu technik doświadczalnych stosowanych w badaniach ludzi. Spotyka się wręcz pogląd, że gdyby nawet wyniki doświadczeń na zwierzętach w ogóle nie mogły być przenoszone na ludzi, to i tak badania te warto byłoby prowadzić po to tylko by doskonalić techniki doświadczalne nauk o zachowaniu ludzkim[222].

[220] Zob. H. F. Harlow, J. P. Gluck, S. J. Suomi, Generalization of behavioral data between nonhuman and human animals, „American Psychologist" 1972, vol. 27, nr 8.

[221] „Szkic taki składa się z mniej lub bardziej nieokreślonego wskazania uznanych za istotne praw i warunków początkowych i po to, aby stać się pełnowartościowym wyjaśnieniem potrzebuje on »wypełnienia«. To zaś wymaga dalszych badań empirycznych; w szkicu zawarta jest propozycja ich kierunku". C. G. Hempel, Aspects of scientific explanation, New York 1965, s. 238.

[222] Pogląd taki wyrażają w cyt. pracy Harlow i jego współpracownicy.

Ta funkcja doświadczeń na zwierzętach opiera się na ich trzech zaletach. Po pierwsze, na zwierzętach nie tylko łatwiej testować hipotezy, ale też łatwiej wypróbowywać nowe sposoby testowania. Po wtóre, względna prostota aranżowanych sytuacji pozwala łatwiej odkrywać prawidłowości rządzące sytuacją samego badania. Po trzecie, prawidłowości owe mogą mieć większą wartość perswazyjną, jeśli stwierdzone zostały w badaniach zwierząt. W szczególnym przypadku mogą być wykorzystane wszystkie trzy zalety. Instruktywne jest w tym punkcie referowane już doświadczenie Rosenthala i Fode'a[223]. Badacze ci zaczęli podejrzewać wpływ oczekiwań eksperymentatora co do wyniku eksperymentu na zachowanie się badanych. Polecali więc dwóm eksperymentatorom wykonać prosty eksperyment na szczurach — wynik potwierdził ich hipotezę. Mogli więc spodziewać się, że zjawisko to wystąpi w jeszcze większym natężeniu w eksperymentach z ludźmi — tutaj możliwości wpływu eksperymentatora na badanych są nieporównanie większe niż w badaniach nad szczurami. Wiele badań potwierdziło ten domysł i zwróciło uwagę na konieczność kontroli pozastymulacyjnego oddziaływania badacza na badanych i w konsekwencji pozwoliło ograniczać liczbę pseudopotwierdzonych hipotez (w tym właśnie kierunku działał wpływ eksperymentatora).

5. Warto tu zanotować jeszcze jedną korzyść płynącą z badań nad zwierzętami, choć w sytuacji tej zwierzęta nie występują jako modele, lecz jako układy porównawcze dla ludzi. Porównanie wyników doświadczeń, w których czynniki kulturowe i wyższe czynności psychiczne są obecne z wynikami doświadczeń, w których są one z założenia wykluczone, pozwala określać zasięg wpływów biologicznych i kulturowych, wyznaczać modyfikacje powodowane tymi ostatnimi, specyfikować wynikające z psychiki i kultury różnice między zachowaniami społecznymi i ludzkimi a zachowaniem zwierzęcym, tym samym zaś poznawać unikalność człowieka jako gatunku.

Jest to punkt wyjścia psychologii porównawczej, nauki dążącej do wyodrębnienia osobliwości gatunków i formułowania międzygatunkowych uogólnień. Dyscyplina ta nie dzieli swej dziedziny badań na człowieka i zwierzęta, ale na człowieka i inne zwierzęta, nie odmawia człowiekowi jego osobliwości, ale uznaje, że czynią one człowieka i kulturę szczególnie interesującymi w perspektywie porównawczej. Ujęte w tym aspekcie doświadczenia na zwierzętach są próbami określenia następstw „jedynej różnicy", którą jest już nie bodziec eksperymentatora, ale odmienność ewolucji. Badacz stwierdza tu tylko wyniki „eksperymentu" wykonanego przez Naturę.

13.4. ZAŁOŻENIA ZASTĘPCZEGO EKSPERYMENTOWANIA NA ZWIERZĘTACH

Wyróżnione wyżej korzyści pozwalają mówić o metodologicznej doniosłości eksperymentu na zwierzętach dla nauk o zachowaniu człowieka. Jeśli korzyści te polegają na inspiracji w stawianiu hipotez, to dla ich przyjęcia podobieństwo zachowań ludzkich i zwierzęcych jest bez znaczenia — jedynie istotne jest niezależne sprawdzenie hipotez na osobnikach ludzkich, aczkolwiek podobieństwo to ma wpływ na prawdopodobieństwo potwierdzenia się hipotez w nowej dziedzinie. Jeśli natomiast, co szczególnie ponętne, wynikom eksperymentów na zwierzętach przyznaje się wartość uzasadniającą dla twierdzeń o ludziach, to konieczne jest do tego przyjęcie stosownych i uzasadnionych założeń modelowych, tutaj — założeń o analogii między modelującymi sytuacjami zwierzęcymi a odwzorowanymi przez nie kompleksami zachowań ludzkich. W tym miejscu wchodzimy na grunt bardzo niepewny, żeby nie powiedzieć — grząski.

[223] R. Rosenthal, K.L. Fode, The effect of experimenter bias one the performance of the albino rat, „Behavioral Science" 1963, vol. 8, nr 2.

Stefan Nowak tak charakteryzuje tego rodzaju sytuacje metodologiczne: „…typ przyjmowanych założeń o naturze zjawisk badanych w sposób istotny rzutuje na zasadność doboru określonej metody do rozwiązywania danego problemu na terenie danego typu zjawisk. Badacz potrafi niekiedy — dzięki wcześniej podjętym badaniom — ocenić zasadność przyjętych założeń. Częściej jednak nie jest w stanie ocenić zasadności założeń ani oszacować granic czy podać Warunków, w jakich założenia są prawdziwe; są one Wówczas mniej lub bardziej hipotetyczne. Badacze różnią się wówczas poglądami na prawdziwość przyjmowanych założeń a w konsekwencji na przydatność tej czy innej metody badawczej do rozwiązywania tego czy innego problemu”[224]. Uprzedzając dalsze wywody, powiemy, że sytuacja, w jakiej znajduje się przenoszenie wyników badań nad zwierzętami na sytuacje ludzkie należy do tych częstszych.

Wśród założeń eksperymentu modelowego na zwierzętach wyróżnimy najpierw przyjmowaną przez badacza lub szkołę koncepcję człowieka, zwaną często „modelem” człowieka. Konstrukt ten, niełatwy do zdefiniowania, obejmuje pewną, niewielką liczbę bezpośrednio nie weryfikowalnych twierdzeń dotyczących najogólniejszych aspektów człowieka: podstawowych wyznaczników jego zachowania, relacji „jednostka — społeczeństwo” czy, co tutaj ważne — relacji „człowiek — zwierzę”.

Rozmaite kierunki w naukach o zachowaniu rozmaicie odpowiadają na pytania o stosunek „człowiek — zwierzę”. Jedne dostrzegają między zachowaniami ludzkimi i zachowaniem zwierzęcym mimo różnic, zasadnicze podobieństwa a drugie — mimo podobieństw, zasadnicze różnice.

I tak, behawioryści nie odmawiają zachowaniu ludzkiemu osobliwości, dostrzegają je w jego rozmaitości, zróżnicowaniu, złożoności, charakterze i ilości sytuacji nagradzających i karzących, roli języka w procesie uczenia się, niemniej jednak utrzymują, że podstawowe prawidłowości zachowania nie muszą być różne, zwłaszcza zaś nie są koniecznie różne pod tymi względami, które behawiorystów szczególnie interesują tj. zachowaniem jawnym i uczeniem się[225].

Drugi biegun zajmują różne odmiany psychologii humanistycznej; weźmy tu za przykład fenomenologiczną koncepcję osoby ludzkiej w sformułowaniu Schelera[226]. Scheier wyróżnia pięć poziomów psychicznego istnienia człowieka w świecie: poziom impulsów emocjonalnych, działających bez udziału świadomości, instynktów, odruchów warunkowych, inteligencji praktycznej — wszystkie one występują i u zwierząt; wreszcie poziom — samoświadomości niezależnej od zdeterminowania organicznego, instynktowego, środowiskowego, właściwy wyłącznie człowiekowi. Przy takiej koncepcji człowieka, badania na zwierzętach byłyby pomocne tylko w zrozumieniu tych poziomów, które są wspólne dla człowieka i dla zwierząt, to jednak, co dlań konstytutywne, wymyka się wszelkim analogiom ze zwierzętami.

Takie, niezwykle ogólne — jak widać — założenia są nie tyle uogólnieniem faktów empirycznych, co na poły filozoficznymi deklaracjami. Mają one niewielką, jeśli w ogóle, wartość uzasadniającą dla ekstrapolacji w zagadnieniach szczegółowych. Nie na tym jednak polega ich rola. One co prawda niczego nie uzasadniają, ale nadają ogólny kierunek i rozmach badaniom człowieka.

Ze stanowiska nauki empirycznej dobrze jest potraktować takie założenia jako dyrektywy heurystyczne, raz nakazujące poszukiwać wspólnych prawidłowości zachowania się zwierząt i ludzi drugi raz — osobliwości zachowań ludzkich i społecznych. Wtedy wybór jednej a nic drugiej możliwości wywodzić trzeba nie z wiedzy o faktach, ale z wartości. Choć pasjonujące może być zarówno wykrywanie podstaw międzygatunkowych uogólnień jak i oznaczanie unikalności człowieka, to w rozmaitych sytuacjach naukowych i kulturowych jedno lub drugie zadanie mo-

[224] S. Nowak, Metodologia badań socjologicznych, Warszawa 1970, s. 249 - 50.

[225] Zob. B. F. Skinner, Science and human behavior, New York 1953, s. 38 - 9; Homans, op. cit., rozdz. 2.

[226] Zob. A. Schutz, Scheier’s theory of intersubjectivity and the general thesis of the Alter Ego, w: Collected papers, vol. 1, The Hague 1962.

że wydać się szczególnie porywające, gdyż np. burzy uznane schematy. Przyjęcie w swoim czasie przez behawioryzm dyrektywy nakazującej szukać praw zachowania wspólnych dla ludzi i zwierząt dało znakomite rezultaty — okazała się więc ona płodną. Ale uświadomienie sobie ograniczeń koncepcji „S—R" spowodowało upowszechnienie się psychodynamicznych, poznawczych modeli człowieka, badania przez nie inspirowane, w których rzadko lub nigdy używano zwierząt, doprowadziły do pogłębienia znajomości specyficznie ludzkich aspektów zachowania.

Równie dogodnie jest dostrzec w różnych ujęciach relacji „człowiek — zwierzę" wyznanie wiary a nie tezy o naturze rzeczywistości. Taka interpretacja zgodna jest z tym, co twierdzili o swoich poglądach sami badacze. Piszą Harlow, Gluck i Suomi: „Istnieje tylko jeden sposób sprawdzenia granic uogólnienia międzygatunkowego a tym jest eksperymentowanie. Pozytywne wyniki mogłyby dowieść istnienia ogólności. Negatywne wyniki nie mogą dowieść niczego oprócz ograniczeń metod lub umysłu eksperymentatora. Stanowisko optymistyczne polega na przekonaniu, że maksymalny stopień uogólnienia nie został jeszcze odkryty. Bezpiecznie jest utrzymywać założenie, że świat obfitości analogii między człowiekiem a zwierzęciem ma dopiero być odkryty. Ludzie i metody będą przypuszczalnie rozwijać się a nie cofać się w rozwoju. Jest to oczywiście wyznanie wiary"[227].

Krokiem dalej w stosunku do założeń tak ogólnych jak wyżej rozważone jest określenie dziedzin, w których — w świetle wiedzy dotychczasowej — istnieją zasadnicze różnice między dziedziną modelującą i modelowaną oraz dziedzin, w których różnice te są małe. Pracę taką wykonali Hebb i Thompson[228]. Ich zdaniem, podstawowe różnice między zwierzętami a człowiekiem istnieją w sferze komunikacji i współpracy, zaś podstawowe podobieństwa w dziedzinie motywacji. Również wg Harlowa i współpracowników motywacja jest tą dziedziną, w której możliwości interesujących nas uogólnień są największe; drugą taką dziedziną jest uczenie się, lecz tylko w prostszych formach, gdyż język i myślenie dają początek zgoła nowym zjawiskom.

Krokiem równie jak poprzedni koniecznym jest określenie granic filogenetycznych analogii. Zważywszy na fakt ogromnego zróżnicowania królestwa zwierząt mówienie po prostu o relacjach „człowiek — zwierzę" jest zwykłym nadużyciem językowym. Istotne różnice istnieją tutaj między zwierzętami niższymi i wyższymi. Podstawowe prawa teorii uczenia występują nawet na bardzo słabo zaawansowanych szczeblach rozwoju ewolucyjnego, ale ciekawsze struktury motywacyjne i operacyjne praktyczne dopiero u małp. Zagadnienie to jest niezwykle trudne, gdyż nie można przyjąć generalnie, że pozycja gatunku na skali filogenetycznej wyznacza stopień podobieństwa jego zachowania się do zachowania się człowieka. Z jednej strony stwierdzono przekonująco brak tej zależności w dziedzinie tzw. obyczajów społecznych[229], z drugiej zaś wyraźną zbieżność między poziomem filogenetycznym a złożonością zachowania w sferze seksualnej[230].

Określać granice filogenetyczne uogólnień międzygatunkowych byłoby lepiej przez charakterystykę struktur organicznych odpowiedzialnych za podobieństwa i różnice w zachowaniu między gatunkami. Myślimy tu nade wszystko o cechach układu nerwowego. Gdyby udało się powiązać prawidłowości zachowania z cechami układu nerwowego, to ze stwierdzenia analogii w zakresie tych cech i założenia nieistotności różnic tych cech dla zachowania można byłoby wnioskować o podobieństwie zachowania. Dużą wartość predyktywną przypisuje się analogiom w budowie mózgu (masa, bogactwo pofałdowań kory mózgowej, stopień komplikacji budowy komórek na zewnętrznym skraju kory mózgowej), choć „nie ma żadnej pojedynczej lub prostej miary anatomicznej, która tłumaczyłaby istnienie bądź brak uogólnień między zwierzętami o względnie różnej pozycji filogenetycznej"[231].

[227] Podobny pogląd wyraża Skinner, The behavior of organisms, op. cit., s. 441 - 2.

[228] D. O. Hebb, W. R. Thompson, The social significance of animal studies, w: G. Lindzey, E. Aronson (eds.), The handbook of social psychology, vol. 2, Reading, Mass. 1968.

[229] Ossowski, op. cit.

[230] Hebb, Thompson, op. cit.

[231] Harlow, Gluck, Suomi, op. cit.

W przypadku idealnym rozwinięte założenia eksperymentu modelowego na zwierzętach powinny więc określać:

a. dziedzinę zjawisk modelowych czyli przedmiot analogii;

b. granice filogenetyczne analogii; dokładniej — powinny przesądzać pozytywnie kwestię istnienia analogii między wykorzystywanym modelem zwierzęcym a ludzkim oryginałem;

c. zjawiska decydujące o tym, że obiekt badany pod danym względem na model się nadaje;

d. powinny też uznawać specyficzne cechy oryginału za nieistotne dla przebiegu zjawisk odwzorowywanych.

Faktycznie natomiast założenia te rzadko są sformułowane, a jeśli już. to w postaci szczątkowej. Formułują je w dodatku raczej- teoretycy i metodologowie niż eksperymentatorzy sami.

Zwykle zakłada się po prostu, że analogia między zachowaniem zwierząt określonych gatunków i zachowaniem ludzi w badanej dziedzinie istnieje, lecz nie podaje się empirycznych jej dowodów ani przyczynowego wyjaśnienia. Uznaje się je — być może — za dobrze znane społeczności naukowej lub też za zbyt trudne do precyzyjnego określenia na tym etapie rozwoju nauki. W rozważaniach z zakresu teorii uczenia pisze się np. tak: „Przyjmujemy, że podstawowe mechanizmy uczenia się i konfliktu u człowieka i innych ssaków są na tyle do siebie podobne, że wyniki badań zwierząt dotyczyć będą i człowieka. Dokładniej, przyjmujemy, że wszystkie procesy psychologiczne stwierdzone u innych ssaków, prawdopodobnie będą obecne również u człowieka”[232]. Albo stwierdza się bardziej szczegółowo, że „chociaż małpki dojrzewają o wiele szybciej, to jednak ogólny wzorzec ich rozwoju jest taki sam, jak u dzieci ludzkich”[233] co uzasadnia wnioskowanie z zachowania pierwszych o zachowaniu drugich.

W innych przypadkach założenie o analogii popiera się wynikami odpowiednich badań porównawczych. Sformułowana w pracy o konfliktach i nerwicach u ludzi i zwierząt teza, że „w sytuacjach konfliktu motywacyjnego ludzie zachowują się w sposób podobny do zachowania się zwierząt w podobnych warunkach”[234], jest podsumowanie przeglądu wyników badań nad reakcjami ludzi i zwierząt na konflikt. Tezę tę uznaje się za coś więcej niż zwykłe uogólnienie sprawozdawcze — nadaje się jej godność prawa nauki. W takiej sytuacji można już przenosić na oryginał nowe zależności stwierdzone w modelu zwierzęcym. Jeśli ponadto dają się one wywieść z analogii już ustalonych, to prawdopodobieństwo ich wystąpienia w nowej dziedzinie jest wysokie. Jeśli zaś ich związek z relacjami znanymi jako wspólne trudny jest do określenia, to i tak mają one dużą wartość jako hipotezy.

Wszędzie tu jednak uderza brak precyzji nieodzownej do nadania wynikom badania zwierząt wysokiej wartości predyktywnej względem zachowań ludzkich. Choć bowiem wiadomo, że między pewnym gatunkiem zwierzęcym a człowiekiem istnieje wspólnota jakichś ogólnie nazywanych procesów, np. uczenia się, to nie wiadomo, które z procesów szczegółowych zaobserwowanych u zwierzęcia występują bez wątpienia i u człowieka. Na to jednak rady nie ma, bo nauka empiryczna, nieaksjomatyczna nie dopuszcza możliwości sformułowania twierdzeń, z których reszta możliwej wiedzy wyniknie logicznie a badacz nie będzie musiał nawet sprawdzać trafności empirycznej tych konsekwencji. Stąd wnioskowanie w modelu zawsze jest obarczone pewnym ryzykiem, choć wtedy, gdy ekstrapolowane ustalenia wynikają logicznie z twierdzeń o których wiadomo, że obowiązują w obydwu układach, wnioskowania mogą być subietywnie pewne. Znaczy to jednak tyle

[232] Miller, Liberalization of basic S-R concepts, op. cit., s. 204.
[233] Harlow, op. cit., s. 110.
[234] R. W. Russel, The comparative study of „conflict” and „experimental neurosis”, „British Journal of Psychology”, Gen. Section 1950, vol. 41, nr 3-4, s. 99.

tylko, że wnioski będą prawdziwe, gdy prawdziwe będą ich przesłanki, uprzednio jasno sformułowane. Nic więcej. I nawet sam demon Laplace'a byłby tu bezsilny.

Wnioskowania z modelu zawsze będą hipotezami, które ostrożniejsi badacze uznają dopiero wtedy, gdy na materiale ludzkim znajdą ich potwierdzenie, a śmielsi uznają natychmiast, jeśli nie znają danych ludzkich z nimi sprzecznych.

13.5. OGRANICZENIA I KIERUNKI KRYTYKI EKSPERYMENTOWANIA ZASTĘPCZEGO NA ZWIERZĘTACH

Mimo wskazywanych wyżej zasług, modelowe eksperymentowanie na zwierzętach budzi także poważne krytyki:

1. Istnieje cały szereg dziedzin, dla wyjaśnienia których wyniki doświadczeń na zwierzętach są z gruntu niewystarczające — są to dziedziny zjawisk specyficzne dla człowieka[235] Przekonać może o tym choćby spis treści podręcznika psychologii społecznej oraz sukcesy kierunków opozycyjnych wobec teorii uczenia się (opartej w znacznym stopniu na badaniach zwierząt), podnoszących rolę czynników poznawczych w wyznaczaniu zachowań ludzkich.

2. Niewątpliwie najsłabszą stroną wykonywanych eksperymentów na zwierzętach jest niedostateczna precyzja zakresu analogii między modelem a oryginałem. Stąd ekstrapolacja wyników doświadczeń ze zwierzętami na dziedzinę zachowań ludzkich a zwłaszcza funkcjonowanie systemów społecznych łatwo wychodzi poza granice dozwolone zasięgiem faktycznej analogii i wywołuje posądzanie o bezpodstawność. Badaczom usiłującym rozciągnąć prawidłowości zachowania się zwierząt w laboratoriach na sytuacje społeczne zarzuca się właśnie brak rozróżnienia tych poziomów analizy, choć nie zaprzecza się, że „jeśli człowieka postawimy w sytuacji analogicznej do gołębia w klatce, to będzie on zachowywał się w podobny sposób"[236]. Jak widać, w krytykach tego typu nie nazywa się uogólnień eksperymentalnych fałszywymi, ale nieadekwatnymi — prawdziwymi w węższej klasie zjawisk niż sądzą ich autorzy.

3. Poznanie człowieka przez badanie zwierząt, zwłaszcza jeśli jest szczególnie preferowane i powszechne, budzi sprzeciw ideowy podjęty w imię obrony humanistycznej wizji człowieka. Doktryna absolutyzująca takie badanie, będąca pewną formą scjentyzmu „zakłada — pisze Schoeck — cyniczny obraz świata w oryginalnym sensie słowa: jest to psiokształtna czy raczej szczurokształtna wizja człowieka. Człowieka rozumie się najlepiej — twierdzi scjentysta — gdy ogląda się go z poziomu gryzonia gorliwie uczącego się zakamarków labiryntu"[237] Perspektywa ta znajdowała wyraz także w zubożaniu języka nauk o człowieku tylko dlatego, że desygnaty wielu jego terminów trudno było znaleźć u zwierząt.

4. Wreszcie wskazuje się, że gwałtowna w swoim czasie ekspansja badań ludzi przez badania zwierząt miała podłoże pozamerytoryczne: nie rzeczywistą, dowiedzioną wartość tych badań dla zrozumienia człowieka, ale łatwość stosowania w nich ścisłych i obiektywnych technik: „Przez badania szczurów, a nie ludzi, zdobywamy status uczonych, gdyż możemy jak przyrodnicy — używać precyzyjnych technik i inter- subiektywnych sposobów przekazywania informacji. To dążenie do statusu jest z naszej strony zrozumiałe, ale z jego powodu stoimy wobec niebezpieczeństwa stracenia dostępu do prawdziwego źródła dostojeństwa dojrzałych nauk. Swej godnej pozazdroszczenia chwały nie zawdzięczają one wierności dla zespołu konwencjonalnych metod, ale raczej bezprzykładnej władzy w przewidywaniu i kontroli biegu natury, którą dały ludzkości dla jej własnego dobra. Psychologia, jako dojrzała nauka, również znajdzie swe usprawiedliwienie nie w wykonywaniu

[235] Wprawdzie w człowieku tkwi cale zwierzę, lecz w zwierzęciu — niecały człowiek (mądrość chińska).
[236] G. Jahoda, Psychologia przesądu, Warszawa 1971, s. 123.
[237] H. Schoeck, Introduction, w: H. Schoeck, J. W. Wiggins (eds.), Scientism and values, Princenton 1960, s. X (określenie „cyniczny" pochodzi od greckiego słowa kynikós = psi).

rytuału metody, ale w przysparzaniu ludzkości władzy potrzebnej do osiągnięcia tych celów" — pisał Allport, wybitny krytyk ekspansji metod eksperymentów zwierzęcych na teren nauk humanistycznych[238].

Wszystkie te zarzuty dotyczą jednak nie samej celowości prowadzenia zastępczych doświadczeń na zwierzętach, ale dotyczą przerostów tej metody. Są one reakcjami obronnymi nauki na nadmierną jej ekspansję. Obydwa zjawiska: hipertrofia jakiegoś podejścia i hamowanie jej przez inne orientacje metodologiczne są często w historii i życiu nauki spotykane, Jest to proces naturalny i nieunikniony, jeśli zważyć, że empirycznych warunków stosowalności metody nie sposób podać *a priori,* a ich uświadomienie przychodzi dopiero wraz z wyczerpaniem się właściwych dla danego stanu wiedzy możliwości poznawczych metody; co więcej — granice te zakreślają raczej krytycy niż zwolennicy tej metodologii.

[238] G. W. Allport, The psychologist's frame of reference, „Psychological Bulletin" 1940, vol. 37, nr 1.

SŁOWO KOŃCOWE

Celem niniejszej pracy nie było w żadnym razie uzasadnienie jakiejś jednej naczelnej tezy. Jednakże teza taka się w niej znajduje. Jest nią przekonanie, że możliwości eksperymentowania w badaniach społecznych są większe, niż sądzi się potocznie i że bariery eksperymentu nie są tak sztywne, jak się wydaje i dałoby się je znacznie przesunąć, gdyby uprzednio przesunąć granice wiedzy merytorycznej i metodologicznej wyobraźni. Rzecz jasna, jest to teza mało — wskutek swego umiarkowania — efektowna i nie warto byłoby dla niej samej zaczynać roboty, którą właśnie kończymy.

Niektóre z barier eksperymentu istotnie są nie do usunięcia, przynajmniej póty, póki chce się eksperymentować bezpośrednio na budzących ciekawość zjawiskach. Myślimy tu przede wszystkim o barierach aksjologicznych i o tych z barier technicznych, które związane są z historycznym charakterem zjawisk społecznych, ich komplikacją i długością trwania. Jednakże badacz zjawisk społecznych, świadom zasobów instrumentarium swych nauk, nie musi wcale doświadczać tej okoliczności jako frustrującej. Obszary zjawisk rozciągające się za linią słupów granicznych z napisami „Nie można!" i „Nie wolno!" mogą być i są z powodzeniem uprawiane innymi narzędziami; eksperyment nie musi nadawać się do wszystkiego. Przed tą linią pozostaje i tak ogromne pole dla inwencji eksperymentatorów.

Badania merytoryczne i metodologiczne mogą ujawnić rozmaite rodzaje zakłóceń manipulacji eksperymentalnych i wskazać lub wynaleźć takie schematy i takie procedury, które pozwolą je kontrolować. Badacze mogą połączyć swe wysiłki z administratorami życia społecznego i przez podejmowanie śmiałych eksperymentów społecznych w ograniczonej skali, przesuwać bariery eksperymentu wywodzące się z ich nikłej władzy nad rzeczami i ludźmi.

Istotna dla oceny wartości wyników badań eksperymentalnych kategoria sztuczności sytuacji eksperymentalnych jest miarą raczej naszej ignorancji niż rzeczy samych. Poczucie sztuczności kurczy się, gdy właściwie pojmuje się istotę eksperymentu, gdy dysponuje się wynikami badań sytuacji społecznej eksperymentu, gdy posiada się wiedzę pozwalającą przechodzić od uproszczonych sytuacji eksperymentalnych do złożonych sytuacji codziennych.

Wiele trudności eksperymentowania, inaczej niepokonalnych, można obejść zastępując trudne do manipulacji sytuacje ludzkie i społeczne ich łatwymi do manipulacji modelami wziętymi z innych poziomów organizacji społecznej lub z królestwa zwierząt. Narzędzie wówczas wydłuża się, ale i staje się wątlejszym zarazem, a jest to znowu zagadnienie posiadania odpowiednich teorii pozwalających ekstrapolować wyniki badań, przynajmniej hipotetycznie, z modelów na oryginały. Wolno sądzić — za wybitnymi eksperymentatorami — że liczne możliwości takich wnioskowań nie zostały jeszcze wykryte.

Zatem, to nie tylko niedostępna, skomplikowana i nienaruszalna pod groźbą stania się sztuczną rzeczywistość zniechęca do eksperymentowania. Rolę tę pełni z powodzeniem również sama nasza niedostatecznie rozwinięta wiedza. A na to rada jest tylko jedna — więcej badań! Również — eksperymentalnych.

NAJWAŻNIEJSZA LITERATURA

Ajdukiewicz K., *Logika pragmatyczna,* Warszawa 1975.

Allport G. W., *Historical background of modern social psychology,* w: G. Lindzey, E. Aronson (eds.), *The handbook of social psychology,* Reading, Mass. 1988, vol. 1.

Bennett C. A., Lumsdaine A. A. (eds.), *Evaluation and experiment: some critical issues in assessing social programs,* New York 1975.

Campbell D. T., *Reforms as experiments,* w: E. L. Struening, M. Guttentag (eds.), *Handbook of evaluation research,* Bewerly Hills 1975.

Campbell D. T., Stanley J. C., *Experimental and quasi-experimental designs for research,* Chicago 1973.

Carlsmith J. M., Eksworth P. C., Aronson E., *Methods of research in social psychology,* Reading, Mass. 1976.

Cook Th. D., Campbell D. T., *The design and conduct of quasi-experiments and true experiments in field settings,* w: M. D. Dunnette (ed.), *Handbook of industrial and organizational psychology,* Chicago 1976.

Cranach M. von (ed.), *Methods of inference from animal to human behavior,* Chicago 1976.

Edwards A. L., *Experimental designs in psychological research,* New York 1968.

Evan W. M. (ed.), *Organizational experiments: laboratory and field research,* New York 1971.

Fairweather G. W., Tornatzky L. G., *Experimental methods for social policy research,* Oxford 1977.

Feibleman J. K. *Scientific method: the hypothetico-experimental laboratory procedure of the physical sciences,* The Hague 1972.

Glass G.V., Wilson V.L., Gottman J.M., *Design and analysis of time-series experiments,* Boulder 1975.

Harlow H. F., Gluck J. P., Suomi S. J., *Generalization of behavioral data between nonhuman and human animals,* „American Psychologist" 1972, vol. 27, nr 8.

Hersen M., Barlow D. H., *Single-case experimental designs: strategies for studying behavior change,* New York 1977.

Hovland C. I., *Reconciling conflicting results derived from experimental and survey studies of attitude change,* w: P. F. Lazarsfeld i in. (eds.), *Continuities in „The language of social research",* New York 1972.

Hovland C. I., Lumsdaine A. A., Sheffield F. D., *Experiments on mass communication,* Princeton 1949.

Kanter R. M., *Commitment and community. Communes and utopias in sociological perspective,* Cambridge, Mass. 1972.

Katz J. D. (ed.), *Experimentation with human beings,* New- York 1972.

Miller A.G. (ed.), *The social psychology of psychological research,* New York 1972.

Oktaba W., *Elementy statystyki matematycznej i metodyka doświadczalnictwa,* Warszawa 1977.

Riecken H. W., Boruch R. F. (eds.), *Social experimentation: a method for planning and evaluating social intervention,* New York 1974.

Rosenthal R., Rosnow R. L. (eds.), *Artifact in behavioral research,* New York 1969.

Rosnow R. L., Robinson E. J. (eds.), *Experiments in persuasion,* New York 1967.

Rossi P. H., Lyall K., *Social policy and social experiments,* New York, Russel Sage Foundation, w druku.

Silverman I., *The human subject in the psychological laboratory,* New York 1977.

Skinner B. F., *Walden Two,* New York 1948, nowe wyd. 1976.

Sułek A. (oprac.), *Logika analizy socjologicznej,* Warszawa, Wyd. UW, w druku.
Swingle P. G., (ed.), *Experiments in social psychology,* New York 1969.
Sztompka P., *O pojęciu modelu w socjologii,* „Studia Socjologiczne" 1968, nr 1.
Środowisko a życie psychiczne. Wybór artykułów z „Scientific American", Warszawa 1969.

EXPERIMENT IN SOCIAL RESEARCH SUMMARY

The book attempts at answers to four questions:

1. What is the specific character of experiment and its value for empirical sciences and in particular for social research?

2. What are the forms of experiment in social research?

3. What are the restrictions set on the application of experiment to the research on social phenomena?

4. How social sciences manage the above mentioned restrictions? The book consists of four parts each giving an answer to succeeding questions.

1

Part I begins with the description of the very concept of experiment. The author defines experiment as a repeatable manipulation consisting in a planned change of some factors of the situation under investigation introduced by the researcher along with his control of other factors made in order to get an answer — in the course of observation — as to the effects of the change. The author discusses the functions of experiment i.e. the superior goals it serves. The author's attention is focused on the verification of causal hypotheses as due to this function experiment is so highly respected. The author gives thus the definition of the causal hypothesis itself and formulates three requirements of the causal proof: ascertainment of the co-occurrence of the events involved, determination of the temporal order and elimination of the alternative to the hypothetical cause explanations of the effect occurrence. The author shows how easily experiment helps to get all these kinds of evidence. Other functions of experiment are shown as well: the heuristic and the diagnostic ones. Particular advantages of experiment are revealed. Besides the facility of verification of the causal hypotheses the author mentions such features as the increase of experience, improvement of the quality of observation and this of the description of phenomena, ingenious linking of theory with empirical data and the especially effective realization of the postulate of intersubjective control of scientific research. Due to these functions and advantages experiment has become one of the most powerful mainsprings of the development of natural sciences. A short outline of the experiment history is given here in order to show the course it has taken before its application in social research.

The analysis of experiment presented in this part of the book is mostly of a generally methodological character — it pertains also to the natural sciences. It may therefore serve as the background against which the specific nature of experiment in social research will appear the more evident.

2

In part II the author gives an exposition of the experimental designs and procedures in social research. He formulates three essential requirements of the validity of experiment used for the verification of the causal hypothesis: theoretical validity, internal and external ones. The exposition of the designs and procedures is meant to show how various designs and procedures assure the above mentioned kinds of validity of experiments.

The author considers in the first place the question of translation of theoretical hypotheses into operational hypotheses formulated in the language of treatments and experimental responses. Dangers of manifold theoretical interpretation of the experiment results are emphasized and various methods assuring valid operationalization of the experimental variable are described as well as those of the effective introduction of treatments to the systems under investigation.

Furthermore the author discusses — following Campbell — extraneous factors which may disturb the effect of the treatment and render thereby the evaluation of its real influence difficult. The author writes here of instability, history, maturation, testing, statistical regression, instrumentation (instrument decay), selection and mortality. The control of these disturbances ensures the internal validity of experiment. Some factors which the external validity of experiment depends on i.e. the generalizability of its results are also considered. Only two of these factors

are analysed in this part, namely the interaction of testing and treatment and the carry over effect.

Next, a broad survey of experimental designs is presented based essentially on Campbell and Stanley's work *Experimental and quasi-experimental designs for research*. These designs are derived from Mill's tradition, they are various modifications of the classical design and usually are designs of one treatment considered as a bivalent one in most cases. Multivariate designs originating from Fisher are discussed as well. They enable research on simultaneous influence of many variables, on influence of multivalent variables and on the concurrent influence of many multivalent variables. Different methods of group selection in experiments are presented. These designs are illustrated by experiments on the change of attitudes and those on the functioning of organizations, both in field and in laboratory conditions.

A separate chapter is devoted to social experiments conceived as controlled tests of social change made on a limited scale. The author discusses here in turn the concept of social experiment itself and this of experimental social innovation, the character of hypotheses tested in research of this kind, problems connected with the introduction, control, measure and evaluation of the effects of the experimental innovations and the possible benefits of social experimentation to practice and science; the difficulties it is supposed to meet with are mentioned as well. The characterization of social experiment is given in such a way as to show that it is a transposition of the ideas and methodology of the laboratory experiment and in field experiment to the area of vital and essential practical problems.

3

Part III is devoted to the answer to the question why experiment in social research is applied on such a small scale in spite of its being such an excellent method? The author points out to three kinds of hindrances standing in the way of experimentation in social research.

Technical hindrances lay in the impossibility of provoking and changing phenomena. Due to the historical character of many social phenomena any manipulation with them is unfeasible just as they do not exist anymore and what is left to the researcher is only the inference based on sources. The involvement of social phenomena in more larger structures causes that quite a number of these phenomena cannot be manipulated with without a concurrent change of whole structures. Research workers, men of science and not those of power, have enough authority only to manipulate with small groups formed in an artificial way at that. Many social processes are of a long-range character and even if a researcher were in a position to provoke their occurrence the observation of their whole course would take too long a time.

Much of the conceivable experiments are forbidden by the prevalent systems of values: ideology, ethics, law. These systems raise three obstacles which experiment must not trespass at all or, when doing so, it has to risk to raise the question of axiological ambiguity. No immediately harmful treatment is allowed as well as no treatment inducing harmful consequences, and the very manipulation with people regardless of consequences is dubious from the moral point of view. The last question is discussed at large by the author who illustrates it by examples of deception experiments. Ideological hindrances in the application of social experiments are also discussed here: some social systems do not allow any such attempts fearing they might disturb their balance.

Another obstacle in the use of experiment is the limited value of the extrapolation of the results of research on experimental situations on the so called real-life situations. Considering the reasons of the different behaviour of the subjects in both types of situations the author is of the opinion that they are to be found in the peculiarities of the psychosocial situation of experimental research, in singularities of certain type of experimental groups, namely the laboratory ones and, last but not least, in the nature of the experimental method itself consisting among others in deliberate simplification of complex real-life situations. The discussion

of the above mentioned problems is based on the results of the American research on psychosocial situation of experiment. The author discusses here the legitimacy of objections raised to the artificial character of situations and results of experiment and he shows what interpretation makes this objection valid and when it is pointless. Most especially he emphasizes the fact that the insignificant value of the effects of experiment is due not to the simplified character of situations under examination but to the lack of theoretical knowledge enabling the transposition of conditions created by the researcher on the so called real ones.

4

The last part of the work is devoted to the methods by means of which social sciences try to overcome the above mentioned obstacles. Three groups of such means may be singled out: the improvement of present experimental techniques and the elaboration of new ones, non-experimental research and substituted experiments on models of systems not open to immediate experiment. The first group has been discussed in several paragraphs of the present work, the second is too far removed from the subject matter of this discussion and so the author centres on the third group — the model experiments. He gives a general methodological characterization of model as such then he applies it to the analysis of the experiments in which the non-manipulated great social structures are replaced by small groups and people out of reach of some manipulations are replaced by any.

In the paragraph regarding model experiments on small groups the author shows in what consists the better accessibility of small groups to the experiment and points out advantages of research of this kind. Such experiments are the source of macrosociological hypotheses and they provide legitimation for such theories, they are also helpful in the formulation of fragments of the general theory of social systems. The author determines the requisites which have to be met in order to assure the validity of extrapolations from the small group level on the great structures level. Furthermore, the author discusses advantages and assumptions of psychosocial experiments on children undertaken with the intention to throw some light on the behaviour of mature individuals and those of the ones made on abnormal individuals aiming at the knowledge of behaviour of normal people.

In the chapter discussing experiments on animals the author shows — just as he has done in the previous chapter — the reasons why it is more easy to experiment on animals than on human beings, the benefits of experiments on animals for sciences of human behaviour, assumptions which nave to be realized in order to get those benefits as well as limitations of such experiments and critical trends.

Concluding, the author expresses the opinion that potentialities of experimentation in social sciences are still unsufficiently taken advantage of. It is not only the inaccessible and complex social reality and the danger of its turning artificial when touched that are discouraging. Just the same part is played with much success by the insufficient knowledge of the subject matter and of methodology. There is only one solution to the problem — more research and experimental research as well.